DOCUMENT RESUME

ED 121 629 SE 020 821

AUTHOR Osborne, Alan R., Ed.
TITLE Investigations in Mathematics Education, Vol. 8 No. 2.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 75
NOTE 71p.
AVAILABLE FROM The Ohio State University, Center for Science and Mathematics Education, 244 Arps Hall, Columbus, Ohio 43210 (Subscription, $6.00, $1.75 single copy)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS *Abstracts; Elementary Secondary Education; *Learning; Logic; *Logical Thinking; Mathematical Concepts; *Mathematics Education; Research; *Research Reviews (Publications)
IDENTIFIERS Piaget (Jean)

ABSTRACT

Eighteen research reports related to mathematics education are abstracted and analyzed. The reports abstracted were selected from seven journals, two books, and a technical report. The majority of the articles reviewed in this volume concern reasoning abilities and mathematical concepts. Logical inference is the subject of four of the studies while development of mathematical concepts is the subject of four others. Three papers are concerned with extension or validation of Piagetian theories. Other reports concern educational games, problem solving, teaching strategies, mastery learning, and aptitude-treatment interaction. Research related to mathematics education which was reported in RIE and CIJE between January and March 1975 is listed. (SD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

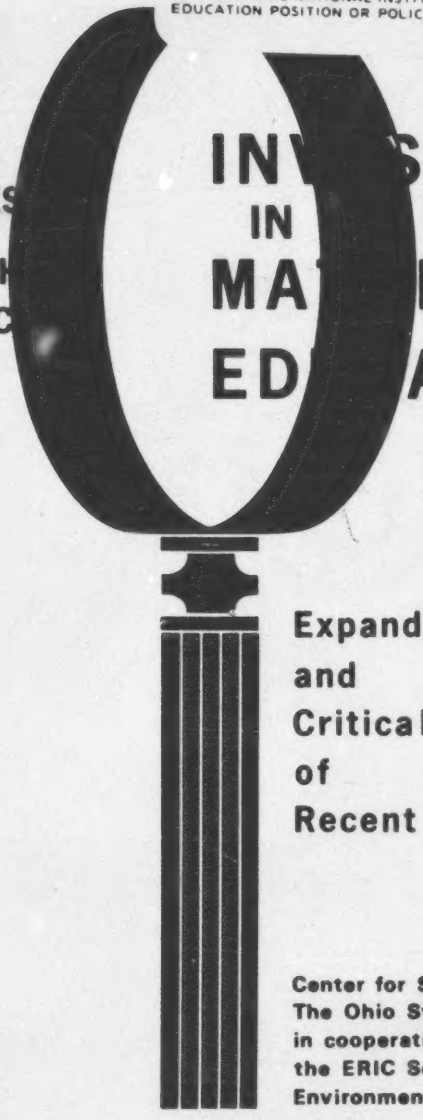
ED121629

INVES
IN
MATH
EDUC

Vol. 8, No. 2 Spring 1975

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY



INVESTIGATIONS IN MATHEMATICS EDUCATION

Expanded Abstracts
and
Critical Analyses
of
Recent Research

Center for Science and Mathematics Education
The Ohio State University
in cooperation with
the ERIC Science, Mathematics and
Environmental Education Clearinghouse

SE020 921

INVESTIGATIONS IN MATHEMATICS EDUCATION

Editor

Alan R. Osborne
The Ohio State University

Associate Editor

Marilyn N. Suydam
The Ohio State University

Advisory Board

E. G. Begle
Stanford University

Joseph N. Payne
University of Michigan

Len Pikaart
Ohio University

J. F. Weaver
University of Wisconsin

Published quarterly by

The Center for Science and Mathematics Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

With the cooperation of the ERIC Science, Mathematics, and Environmental Education Information Analysis Center

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

Mathematics Education Research Studies Reported in Research in Education (January - March, 1975) 1

Mathematics Education Research Studies Reported in Journals as Indexed by CIJE (January - March, 1975) 3

Bass, Helen. Topological Understandings of Young Children. In Rosskopf, M. F. (ed.) Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.
Abstracted by J. LARRY MARTIN 5

Bowers, Robert Garth. The Effect of Triangle-Pair Configuration Variation on Achievement of Selected Classes of Instructional Objectives in Plane Geometry. Journal for Research in Mathematics Education, v4 n2, pp110-113, March 1973.
Abstracted by NORMAN L. WEBB. 9

Brainerd, Charles J. The Origins of Number Concepts. Scientific American, v228 n3, pp101-109, March 1973.
Abstracted by JANE O. SWAFFORD. 13

Carroll, C. A. Low Achiever's Understanding of Logical Inference Forms. In Rosskopf, M. F. (ed.) Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.
Abstracted by LARS C. JANSSON 16

DeVries, D. L.; Edwards, K. J. Learning Games and Student Teams: Their Effects on Classroom Process. American Educational Research Journal, v10, pp307-318, Fall 1973.
Abstracted by ROBERT KALIN. 19

Gregory, John W.; Osborne, Alan R. Logical Reasoning Ability and Teacher Verbal Behavior Within the Mathematics Classroom. Journal for Research in Mathematics Education, v6 n1, pp26-36, January 1975.
Abstracted by THOMAS J. COONEY. 23

Jerman, Max. Individualized Instruction in Problem Solving in Elementary School Mathematics. Journal for Research in Mathematics Education, v4 n1, pp6-19, January 1973.
Abstracted by GERALD A. GOLDIN. 27

Johnson, David C. Learning of Selected Parts of a Boolean Algebra by Young Children. In Steffe, L. P. (ed.) Research on Mathematical Thinking of Young Children: Six Empirical Studies. Reston, Virginia: NCTM, 1975.
Abstracted by JOHN R. KOLB. 31

Just, M. A.; Carpenter, P. A. Comprehension of Negation and Quantification. Journal of Verbal Learning and Verbal Behavior, v10, pp244-253, 1971.
Abstracted by ROBERT B. KANE. 34

Miller, Scott A. Contradiction, Surprise, and Cognitive Change: The Effects of Disconfirmation of Belief on Conservers and Nonconservers. Journal of Experimental Child Psychology, v15 n1, pp47-62, February 1973.
Abstracted by J. DAN KNIFONG. 37

Montgomery, Mary E. The Interaction of Three Levels of Aptitude Determined by a Teach-Test Procedure with Two Treatments Related to Area. Journal for Research in Mathematics Education, v4 n4, pp271-278, November 1973.
Abstracted by AARON D. BUCHANAN 41

Phillips, Robert B., Jr. Teacher Attitude as Related to Student Attitude and Achievement in Elementary School Mathematics. School Science and Mathematics, v73 n6, pp501-507, June 1973.
Abstracted by RICHARD HUNKLER 45

Romberg, Thomas A.; Shepler, Jack. Retention of Probability Concepts: A Pilot Study Into the Effects of Mastery Learning with Sixth-Grade Students. Journal for Research in Mathematics Education, v4 n1, pp26-32, January 1973.
Abstracted by PAUL R. TRAFTON 48

Taback, S. The Child's Concept of Limit. In Rosskopf, M. F. (ed.), Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.
Abstracted by JOHN G. HARVEY. 51

Taloumis, Thalia. The Relationship of Area Conservation to Area Measurement as Affected by Sequence of Presentation of Piagetian Area Tasks to Boys and Girls in Grades One Through Three. Journal for Research in Mathematics Education, v6 n4, pp232-242, November 1975.
Abstracted by LESLIE P. STEFFE. 55

Wearne, Diana; Whitaker, Donald; Romberg, Thomas A. Study II--An Investigation of Three Different Response Formats for a Two-Digit Addition and Subtraction Test. In Romberg, Thomas A. (ed.), Four Studies on Aspects of Assessing Computational Performance. Technical Report No. 297. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin-Madison, July 1975.
Abstracted by BARBARA J. PENCE. 58

Webb, Leland F.; Carry, L. Ray. Interaction of Spatial Visualization and General Reasoning Abilities with Instructional Treatment in Quadratic Inequalities: A Follow-Up Study. Journal for Research in Mathematics Education, v6 n3, pp132-140, May 1975.
Abstracted by CLYDE A. WILES. 61

Weitz, Lawrence J.; Bynum, Terrell Ward; Thomas, James A.; Steger, Joseph A. Piaget's System of 16 Binary Operations: An Empirical Investigation. The Journal of Genetic Psychology, v123, pp279-284, 1973.
Abstracted by SUZANNE K. DAMIRIN and RICHARD J. SHUMWAY . . .64

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESEARCH IN EDUCATION
January - March 1975

ED 095 829 Fisher, Maurice D.; And Others. Student Control and Choice: Their Effects on Student Engagement in a CAI Arithmetic Task in a Low-Income School. Technical Report No. 41. 30 p. MF and HC available from EDRS.

ED 095 871 Patterson, Eldon. An Evaluation of Project PLAN. 450p. MF and HC available from EDRS

ED 096 100 Layton, David, Comp. A List of Theses for Higher Degrees in British Universities Completed 1968, 1969, 1970, and 1971. 24 p. MF and HC available from EDRS.

ED 096 160 Eisenberg, Theodore A. An Analysis of Computational Errors Made by Teachers of Arithmetic: 1930, 1973. 16p. MF and HC available from EDRS.

ED 096 334 Evans, Charles L. A Comparison of "Continuous Progress" and "Traditional" Classes in Three Middle Schools. 103p. MF and HC available from EDRS.

ED 097 049 Moore, Claude. A Study of the Effectiveness of the Developmental Mathematics Courses at Danville Community College. 20p. MF and HC available from EDRS.

ED 097 205 Groen, Guy J. Basic Processes in Simple Problem Solving. Final Report. 16p. MF and HC available from EDRS.

ED 097 225 Suydam, Marilyn N. A Categorized Listing of Research on Mathematics Education (K-12) 1964-1973. 364p. MF and HC available from EDRS.

ED 097 229 Bell, Marilee; Starkey, John D. The Relationship Between Parent's Education and Performance on Math and Reading Tests. 10p. MF and HC available from EDRS.

ED 097 231 Gray, T. Harrison; And Others. Speed and Accuracy of Addition in Normal Time and Decimal Time Systems. 37p. MF and HC available from EDRS.

ED 097 290 Schall, William E.; And Others. Developing Mathematical Processes (DMP). Field Test Evaluation, 1972-1973. 49p. MF and HC available from EDRS.

ED 097 910 Netusil, Anton J.; Kockler, Lois H. CAI: Overcoming Attitude Barriers. 10p. MF and HC available from EDRS.

ED 098 048 Hollingsworth, Caroline Dean. A Comparison of Verbal and Nonverbal Instruction in Elementary School Mathematics. 150p. Not available from EDRS. Available from University Microfilms (74-6594).

ED 098 049 Burke, Gerald Clayton. Case Studies of High School Students Using Physical Models to Study Mathematical Systems. 74p. Not available from EDRS. Available from University Microfilms (74-6013).

ED 098 050 Harutunian, Harold. Validation of a Learning Hierarchy Using Classroom Interaction. 173p. Not available from EDRS. Available from University Microfilms (74-7621).

ED 098 051 Ladd, Norman Elmer. The Effects of Electric Calculators on Attitude and Achievement of Ninth Grade Low Achievers in Mathematics. 126p. Not available from EDRS. Available from University Microfilms (74-6225).

ED 098 052 Roderick, Stephen Alan. A Comparative Study of Mathematics Achievement by Sixth Graders and Eighth Graders, 1936 to 1973, 1951-55 to 1973, and 1965 to 1973. 373p. Not available from EDRS. Available from University Microfilms (74-7423).

ED 098 053 Brown, John Kenneth, Jr. Textbook Use by Teachers and Students of Geometry and Second-Year Algebra. 177p. Not available from EDRS. Available from University Microfilms (74-5534).

ED 098 054 Kidder, F. Richard. An Investigation of Nine-, Eleven-, and Thirteen-Year-Old Children's comprehension of Euclidean Transformations. 25p. MF and HC available from EDRS.

ED 098 081 Pereira, Waldecyr Cavalcanti De Araujo. The Relationship Between Girls' Overt Verbalization, Performance, Retention, Rules and Strategies as They Learn a Mathematical Structure: A Study Based on Elements of a Potential Theory Which Relates Thinking, Language and Learning. Technical Report No. 293. 170p. MF and HC available from EDRS.

ED 098 254 Babikian, Elijah; Buchanan, Aaron. Developing a System of Criterion Referenced Assessment-Reteaching Cycles in Textbook Supported Mathematics Instruction. 22p. MF and HC available from EDRS.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION
January - March 1975

EJ 104 714 Gyr, J. W.; And Others. "Do Mathematical Group Invariants Characterize the Perceptual Schema of Younger and Older Children?" Human Development, v17 n3, pp176-186, 74.

EJ 105 676 Engin, Ann W. "The Relative Importance of the Subtests of the Metropolitan Readiness Test in the Prediction of First Grade Reading and Arithmetic Achievement Criteria." Journal of Psychology, v88, pp289-298, Nov 74.

EJ 105 913 Goodstein, H. A.; Kahn, H. "Pattern of Achievement Among Children With Learning Difficulties." Exceptional Children, v41 n1, pp47-49, Sep 74.

EJ 105 960 Thibodeau, Gerard P.; Vitello, Stanley J. "The Use of Programmed Instruction in Developing the Computational Skills of Mentally Retarded Children." Journal of Mental Retardation, v12 n4, pp46-48, Aug 74.

EJ 106 359 Kerslake, Daphne. "Some Childrens Views on Probability." Mathematics in School, v3 n4, pp22, Jul 74.

EJ 106 368 Bailey, Terry G. "Linear Measurement in the Elementary School." Arithmetic Teacher, v21 n6, pp520-525, Oct 74.

EJ 106 393 Karplus, Elizabeth F.; And Others. "Intellectual Development Beyond Elementary School IV: Ratio, The Influence of Cognitive Style." School Science and Mathematics, v74 n6, pp476-482, Oct 74.

EJ 106 400 Okey, James R. "Altering Teacher and Pupil Behavior With Mastery Teaching." School Science and Mathematics, v74 n6, pp530-535, Oct 74.

EJ 106 468 Roberts, A. M. "Modern Versus Traditional Mathematics." Australian Mathematics Teacher, v30 n4, pp139-148, Aug 74.

EJ 106 492 Wittrock, M. C. "A Generative Model of Mathematics Learning." Journal for Research in Mathematics Education, v5 n4, pp181-196, Nov 74.

EJ 106 493 Shumway, Richard J. "Negative Instances in Mathematical Concept Acquisition: Transfer Effects Between the Concepts of Commutativity and Associativity." Journal for Research in Mathematics Education, v5 n4, pp197-212, Nov 74.

EJ 106 494 Dahlke, Richard M. "Determining the Best Predictors of Success and of Time of Completion or Dropout in an Individualized Course in Arithmetic at a Community College." Journal for Research in Mathematics Education, v5 n4, pp213-223, Nov 74.

EJ 106 495 Eisenberg, Theodore A.; McGinty, Robert L. "On Comparing Error Patterns and the Effect of Maturation in a Unit on Sentential Logic." Journal for Research in Mathematics Education, v5 n4, pp225-237, Nov 74.

EJ 106 496 Suydam, Marilyn N.; Weaver, J. F. "Research on Mathematics Education (K-12) Reported in 1973." Journal for Research in Mathematics Education, v5 n4, pp238-272, Nov 74.

EJ 106 607 Arter, Judith A.; Clinton, LeRoy. "Time and Error Consequences of Irrelevant Data and Question Placement in Arithmetic Work Problems II: Fourth Graders." Journal of Educational Research, v68 n1, pp28-31, Sep 74.

TOPOLOGICAL UNDERSTANDINGS OF YOUNG CHILDREN. Bass, Helen. In Rosskopf, M. F. (ed.) Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by J. Larry Martin, Missouri Southern College.

1. Purpose

Tasks were designed to determine whether or not school grade level and socioeconomic climate affect children's understanding of selected topological concepts (order, enclosure, equivalence). The author hypothesized that older children would perform better than younger children and that children from a higher socioeconomic climate would perform better than children from a lower socioeconomic climate.

2. Rationale

Piaget's theory that topological concepts precede projective and Euclidean concepts in the spatial representation of the child underlies the study. An overview of Piaget's topological tasks is followed by brief reports of relevant studies. Although some variance from Piaget's findings exists, replication studies in general tend to corroborate Piaget's assertions. The present study focuses on the child's conception of the topological concepts of order, enclosure, and equivalence.

3. Research Design and Procedure

A total of 180 kindergarten, first-, and second-grade children were randomly selected from two eastern city schools, 30 from each grade at each school. School A was described as middle to upper-middle class in socioeconomic level; school B as middle to lower-middle class. The investigator tested each child individually, using tasks intended to measure understanding of the topological concepts of order, enclosure, and equivalence. All tasks were administered at one sitting of 15-20 minutes.

Each of the four order tasks required the subject to match the color sequence on a model card by selecting from a collection of choice cards. Materials consisted of cards upon which colored disks had been pasted in linear sequence. Task one consisted of five model, or sample, cards each showing a two-disk sequence. They were presented one at a time with the investigator asking the subject, "Where is this one?" or "Now find this card," from between two choice cards. Tasks two, three, and four also had five model cards each, but the length of the sequence was increased to three, five, and seven disks, respectively, and the subject selected from three, rather than two, choice cards each time. Each child was assigned a score indicating his number of correct responses (20 possible). Brief warm-up tasks were used prior to the first task to insure that the child was capable of distinguishing colors and to familiarize him with the procedures.

The same matching-from-sample technique was used for the equivalence tasks. There were a total of 24 sample cards and for each sample card three choices were available. In each case, one choice card was topologically equivalent to the sample but none was identical to the model. Again, the child's score indicated number of correct responses (24 possible).

Two orientation activities preceded the enclosure tasks--one to determine whether or not the subject understood the difference between open and closed curves, the other to determine whether or not the subject understood the difference between the inside and the outside of the simple closed curve.

Three tasks followed the orientation activities. In the first, the subject was given a loop of yarn and cut-outs of cats and dogs, and asked to fix the fence so that neither animal could run away and so that the dog could not chase the cat. If the subject failed this task, the last two tasks were omitted. Successful subjects went to the second task and were given a third animal, a mouse cut-out, and asked to find a place where it could be safe from both the dog and the cat. For the third task the subject was to place a fourth animal cut-out, a chicken, where it would be safe from the other three animals. Subjects received one point for each correct response (3 possible).

Median scores were computed for each concept. A score was considered "successful" if it was greater than or equal to the median. Numbers of successful students were contrasted by grade and by school for each concept. To determine whether differences were significant, confidence intervals of 0.95 were computed for the differences using Goodman's method (1964).

4. Findings

Concept of Order: Contrasts for which the confidence intervals were significant were:

(a) for each school, kindergarten against second grade (second grade was higher);

(b) for each school, the average of kindergarten and first grade against second grade (second grade was higher);

(c) for school B, kindergarten against the average of first and second grade (average of first and second grade was higher);

(d) for schools combined, kindergarten against first grade (first grade was higher);

(e) for all grades combined, school A against school B (school A was higher);

(f) for kindergarten and second grade combined, school A against school B (school A was higher).

Concept of Enclosure: All children successfully performed both orientation activities. Eighty-four percent of the children in school A performed all three tasks. Seventy percent of the children in school B performed all three tasks. When the classes of the two schools were combined, second grade was significantly higher than kindergarten. There were no significant results pertaining to socioeconomic influence.

Concept of Equivalence: There were no significant results pertaining to grade level. When the schools were compared as a whole, school A showed significantly better results than school B.

5. Interpretations

Results for the concepts of order and enclosure were consistent with Piaget's findings. An understanding of order and enclosure is present in many children at age five and tends to mature around age seven. The year between ages six and seven is especially important. An understanding of equivalence does not mature as early. Maturity for this concept occurs closer to the age of eight. Socioeconomic climate appears to have some influence on the development of all three concepts.

In general, both hypotheses of the study were substantiated. In addition, the concepts studied seem to develop in the sequence: enclosure, order, equivalence.

Critical Commentary

The study does not address itself to the validity of Piaget's hypothesis that topological concepts develop before Euclidean or projective concepts in the representational space of the child. Presumedly, the hypothesis is accepted, although it is pointed out that Lovell suggests it should remain an open question. Many mathematics educators agree with Lovell's suggestion, for a variety of reasons. Perhaps it is best not to debate the validity of Piaget's hypothesis, but merely to accept the child's understanding of topological concepts as worthy of study regardless of Piaget's hypothesis.

To study the child's conception of space requires that a distinction be made between perception and representation. There is no indication that the experimenter made this distinction. Indeed, the matching-from-sample technique seems to measure perceptual rather than representational space. Tasks where the subject must transform or construct are more likely to measure representation.

It is not surprising to find older children performing better than younger, nor advantaged children better than disadvantaged. Nor is it surprising that equivalence should develop later than order and enclosure. Equivalence requires keeping all topological properties invariant, thus would logically require concepts of order and equivalence. It is interesting, however, that the period of transition from the pre-operational to the concrete operational cognitive levels was once again found critical.

A further note on equivalence: Many of the topological equivalents could be found only by performing Euclidean transformations, an ability developed later according to Piaget. Also, the possibility of using the process of elimination exists on the equivalence test, especially if the items were not presented randomly. In many cases the same set of choice cards was used for more than one sample card, with each choice card being correct exactly once. Evidently the subjects were told when they had selected correctly, at least at the beginning of the task.

Some other questions:

(1) Why were 31 contrasts made for each concept, rather than using analysis of variance?

(2) What were the reliabilities for the tests?

(3) Had any of the concepts been included in the formal education of any of the subjects?

(4) Which items were more difficult? An item analysis and, as the experimenter points out, a discussion of the wrong selections would have been helpful.

J. Larry Martin
Missouri Southern College

EJ 075 351 SE 508 025

THE EFFECT OF TRIANGLE-PAIR CONFIGURATION VARIATION ON ACHIEVEMENT OF SELECTED CLASSES OF INSTRUCTIONAL OBJECTIVES IN PLANE GEOMETRY. Bowers, Robert Garth. Journal for Research in Mathematics Education, v4 n2, pp110-113, March 1973.

Expanded Abstract and Analysis Prepared Expecially for I.M.E. by Norman L. Webb, Indiana University.

1. Purpose

To study whether instruction to achieve a set of objectives expressed by using more than one class of congruent triangle-pair referents transfers to:

1. the achievement of the same set of objectives expressed by using classes of congruent triangle-pair referents different from those used in instruction; and

2. the achievement of a different set of objectives over the same classes of congruent triangle-pair referents used in instruction.

2. Rationale

A series of studies was conducted under the auspices of an instructional research project referred to as the Paradigms Project. The overall purpose of the studies was to provide information that could be used to make rational decisions in curriculum construction. Seven of the studies conducted were classified, on the basis of the variables that the studies dealt with, into three subclasses: modes of representation, structure of curriculum hierarchies, and choice behavior.

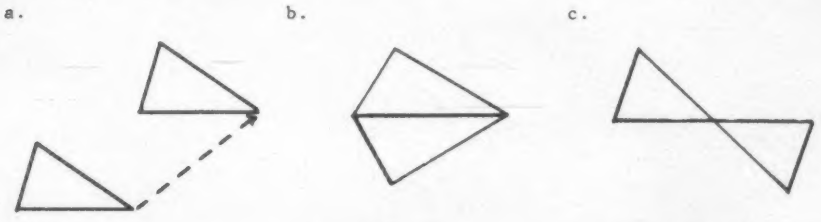
This study is concerned with perceptual problems encountered in teaching the concept of congruent triangles, one type of problem within the domain of the modes-of-representation subclass. The study of such perceptual problems can be systematized by the use of a two-dimensional scheme, position X transformation, used to identify classes of congruent triangle-pairs. The positions of the two triangles can be described by the number of points the triangles have in common (none, a finite number, or an infinite number) and whether or not the interiors of the triangles overlap. The four possible transformations of one triangle of a congruent pair to the other are translation, reflection, rotation, and translation-reflection.

3. Research Design and Procedure

Two sets of objectives were involved in the study. One set of three objectives dealt with the concept of one-to-one correspondence between congruent triangles. A second set of four objectives dealt with the congruency (SAS, ASA, SSS) postulates. Two instructional treatments were identified:

1. A Basic Single-Configuration Instructional Treatment - Instructional sequences to achieve each objective using only one configuration class (See Figure 1a) for the triangle-pair referents.

2. A Multiconfiguration Instructional Treatment - Instructional sequences to achieve each objective using an aggregate of four configuration classes. The classes of triangle-pair referents include the same class as used in the single-configuration instruction plus three more classes: reflected triangles with infinite intersection without overlapping interiors (Figure 1b); reflected triangles with infinite intersection and overlapping interiors; and rotated triangles with finite intersection without overlapping interiors (Figure 1c).



Translation with Empty Intersection	Reflection with Infinite Intersection Without Overlapping Interiors	Rotation with Finite Intersection Without Overlapping Interiors

Figure 1. Examples of Configuration Classes of Congruent Triangles

The hypotheses were tested over a criterion configuration class containing five configuration classes not used in the two instructional treatments. These classes included rotated, reflected, and translated-reflected congruent triangle-pairs with various degrees of overlap between the triangles.

Two samples of students were randomly selected from those students whose performance on a pretest indicated a lack of attainment of the two sets of objectives. Sample S_1 included 41 students from four ninth-grade classes studying geometry in an accelerated program. Sample S_2 included 91 students from four regular tenth-grade geometry classes from a school district different from that of sample S_1. Students in each sample were randomly assigned to one of six experimental treatment groups defined by varying combinations of instructional treatments applied to the two sets of objectives. The two samples of students were judged to be nonhomogenous groups and separate data analyses were performed for each group.

Modules of instruction for each objective over the specified configuration class were mediated by an IBM 1500 Instructional System. All data were collected from on-line instructional treatments.

The Mann-Whitney U Test (α=.05) was used to test the hypothesis that the multiconfiguration treatment will produce greater pre-posttest gain score than the single-configuration treatment over the criterion configuration class for each objective. For hypotheses with only one instructional treatment specified, a control group which received only a basic single-configuration treatment for all objectives pertinent to the hypotheses was used for comparison. Pass-fail classification of objective achievement was used to compare the two treatments. The Fisher and chi square tests (α=.05) were used to test significant differences between treatment and control groups for the testable instances of each hypothesis.

4. Findings

For the four hypotheses that dealt with the effects of various multiconfiguration treatments of objective sets on the achievement, without instruction, of these as well as other objectives over the criterion configuration class, 17 out of 17 for S_1 and 14 out of 17 for S_2 of the testable instances were not supported. For the two hypotheses that dealt with the effects of various multiconfiguration treatments of objective sets on the achievement, without instruction, of different objectives over the configuration classes within the multiconfiguration treatment, none of the 11 test instances for either sample was supported. All four of the testable instances for both S_1 and S_2 were supported for the hypothesis that changing the order of the first two objectives in a set of three objectives will not affect achievement without instruction of a discrimination objective for the set of objectives over the instructional configuration classes.

5. Interpretations

(a) Attributes of both treatments, single and multi-confuguration classes, are highly transferable, without instruction, to like and different objectives over the criterion class of configuration.

(b) There were no statistical differences between the achievement pattern of the treatment groups over the configuration classes used in the multiconfiguration treatment.

(c) Instructional presentation order among the congruency postulates (SAS, ASA, SSS) appears to be unimportant, regardless of the instructional treatment applied, as ascertained by later achievement comparisons.

Critical Commentary

Approaching decision making in curriculum construction in a systematic way through a series of integrated studies is a much needed and admirable task. The scope and the results of this study must be considered with the knowledge that it is one of a series of studies, some of which are attacking similar questions.

Some questions do arise from the report of this study:

1. What was an example of the objectives used in the study?

2. What data supported the conclusion that attributes of both single and multiconfiguration treatments are highly transferable?

3. What were the criteria used to assess the attainment of an objective?

Questions which arise about the conceptualization of the study and suggest directions for include:

1. What was the cognitive level of the objectives used? If the objectives were at the level of knowledge of fact, how would the results vary if higher-level objectives were used?

2. Were there any aptitude-treatment interactions? Little information was provided about either sample group. With such perceptual tasks as used in this study, spatial abilities may have a strong effect on the transfer of instruction.

3. What was the influence of the context in which the objectives were presented? Apparently, the effects of single- and multiconfiguration instruction were the same in achieving objectives presented in the same mode as the instruction. But what if the context were changed, as, for example, by embedding the use of the objective in the process of solving or interpreting a proof? Is there still no difference between having experienced only one triangle-pair referent compared to several triangle-pair referents?

Norman L. Webb
Indiana University

THE ORIGINS OF NUMBER CONCEPTS. Brainerd, Charles J. Scientific American, v228 n3, pp101-109, March 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jane O. Swafford, Northern Michigan University

1. Purpose

The purposes of the three studies described are: (1) to investigate the relationship between the emergence of the concept of ordination and the concept of cardination in young children, (2) to investigate the relation between the emergence of each of these two concepts and competence in natural number computation, and (3) to investigate the cognitive dependence among these concepts.

2. Rationale

There are two different mathematical theories for developing the natural numbers. One, owing to Frege and Russell, is called here the cardination theory. The other, called the ordination theory, is based on the five axioms of Peano. In the cardination theory each natural number n is regarded as the collection ("superordinate class") of all sets ("subordinate classes") which contain precisely n elements. In Peano's theory natural numbers are placed in a transition, asymmetrical relation. Neither theory is mathematically more acceptable than the other. Hence, these studies were designed to investigate whether the actual emergence of numerical ideas in young children supports one of these two mathematical theories. Do children's number concepts develop via cardination or ordination? No previous research is cited.

3. Research Design and Procedure

Subjects for the first study were 180 Canadian children between the ages of 5 and 7. A two-part ordination test consisted of presenting each subject with two clay balls and two dowels to be ordered with respect to weight or length by comparing solely to a third middle weight or middle length ball or dowel. The cardination test consisted of 6 cards containing two parallel rows of between 6 and 10 dots. Each subject was asked to compare the cardinality of the two rows on each card without counting. On both the ordination test and cardination test, the subjects were grouped according to their performance into Stage I, totally incapable of quantification; Stage II, capable of quantification only for "left less than right" in ordination or for "many-to-one" relations in cardination; or Stage III, capable of both kinds of quantification.

The second study involved 90 kindergarten and 90 first-grade subjects. In addition to the ordination and cardination test described above, the first-grade students were given an addition facts test and a subtraction facts test. The kindergarten students were given only the addition test. Students were again sorted into three stages with respect to the ordination and the cardination tests. Students were additionally sorted into

three levels of numerical skill. As with the first study, comparisons were made between the number of children at each level on the ordination and cardination test. Further analysis was made of the composition of each level with respect to the levels of performance on the other tests.

The third study involved 240 subjects between the ages of 5 and 6 divided into four matched groups (2 control, 2 experimental) on the basis of performance on the three tests described above. Over an 8-week period, one experimental group was given corrective feedback during a weekly cardination test and the other experimental group received corrective feedback on a ordination test. The control groups were given weekly tests without feedback. The original ordination, cardination, and computation tests were then readministered. Pre- and post-test results by group were given. Statistical analyses of the results, although used, were not reported.

4. Findings

In the first study, 24, 63, 93 of the subjects were functioning respectively at Stages I, II and III on the ordination test, whereas only 92, 73, 15 were at similar levels on the cardination test. Of the 93 subjects at Stage III in ordination, only 12 were at Stage III in cardination. However, of the 92 subjects at Stage I on cardination, 72 were at Stage II or III on ordination. Similar results were reported for the second study.

Of the 119 students at the top level on ordination on the second study, 71% were average or above on number competence. Of the 95 subjects with no cardination (Stage I), 47% showed average or above number competence.

The statistical analysis of the third study led to the following findings: (1) ordination and cardination improves with feedback, (2) the average improvement in ordination was much better than the average improvement in cardination, (3) the number performance of the control group trained in ordination was superior to the matched control group, and (4) the number performance of the experimental group trained in cardination was not significantly superior to the matched control group.

5. Interpretations

From the first study it was concluded that the concept of ordination emerges in a child's thinking long before the concept of cardination. From the second study, it was concluded that first ordination emerges followed by number competence, and finally by cardination. From the third study it was concluded that the concept of natural number is dependent on a prior understanding of ordination and not on a prior understanding of cardination. The implication drawn from the three studies is that the concepts of ordination should first be introduced to children, with natural being introduced as a by-product of transitive, asymmetrical relations and, finally, with cardination being introduced as a generalization of the natural numbers.

Critical Commentary

The investigators in these studies were victims of the same pitfall attributed to other investigators in developmental psychology. Namely, they are guilty of using "loosely defined methods of assessment that bear only a vague or intuitive resemblance to the (mathematical) concepts that are to be measured." Six cards containing parallel rows of red and blue dots is no more an embodiment of the theory of Ferge and Russell than 3 clay balls and 3 dowels is a reasonable representation of the Peano Axioms. The ordination test is, in fact, a test of transitivity in two specific order relations neither of which is defined on the set of natural numbers. The cardination test is no more than a spatial perception test inasmuch as the subjects were not allowed to match the two rows of dots either through moving the dots or drawing lines. The computation tests, while called at various points in the article measures of natural number concepts and measures of natural number competence, are only computation tests. Inferences about the origin of a child's number concepts can only be drawn from these studies with courage.

The interpretations given are without foundation. It is alarming that the author was allowed the forum of a popular magazine to fault "new math" and current textbooks for using a "cardination" approach to natural numbers as opposed to an "ordination" approach. Since counting, the author warns, is the commonest example of ordination, the characterization of the "new math" treatment of numbers as cardination is misrepresentative if not completely false. What conclusions, if any, that can be drawn from these studies are obscured by tests of the obvious masquerading as tests of the much more complex.

Jane O. Swafford
Northern Michigan University

LOW ACHIEVER'S UNDERSTANDING OF LOGICAL INFERENCE FORMS. Carroll, C. A. In Rosskopf, M. F. (ed.) Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Lars C. Jansson, University of Manitoba.

1. Purpose

The purpose of the study, as stated in the Results section, was (a) "to examine the effectiveness of instructional sessions with small groups as a means of improving performance in conditional reasoning, and (b) to analyze the difficulties that low achievers experience in conditional reasoning."

2. Rationale

Conditional reasoning is placed in a general Piagetian developmental context. Abilities to handle specific forms of valid (modus ponens, contrapositive) and invalid (converse, inverse) arguments are related to the research of Hill (1961), O'Brien & Shapiro (1968), and the Cornell Critical Thinking Project, Phase 1, of Ennis & Paulus (1965). Content of items is related to Piaget's distinction between formal and concrete stages of reasoning and the work of Ennis and Wilkins (1928). Four types of content are employed: concrete familiar, symbolic, misleading, and removed from reality.

3. Research Design and Procedure

Tests of 48 items were validated and shown reliable on a group of average and above-average ability students. Four forms of argument and four content types provided three items in each form/content category. Tests were administered to six groups as follows:

(a) Experimental - logic instruction: 24 boys and 24 girls

(b) Alternate - probability instructions: 12 boys and 12 girls

(c) Contrast - no instruction: 82 boys and 37 girls

Boys and girls groups were in two different schools. Subjects fit the definition given of "low achiever", but assignment was not truly random and was dependent on such matters as student schedules.

Data are reported as percentages of students in each of six groups showing improvement from pretest to posttest. Analysis used was chi square. Pretest-posttest means are reported for form, content, and total, but no tests of significance are reported on these. Graphical displays are employed to discuss pretest-posttest differences for form, content, and their interactions.

4. Findings

a. "H_1: There will be more improvement among students in the experimental group than among those in the control group." Fifteen different comparisons all yielded nonsignificant ($p = .01$) chi square statistics, thus H_1 was rejected.

b. When analyzed by argument form, only the converse for the experimental groups showed significant gains. Trends favored experimental groups for modus ponens and inverse, control groups on the contrapositive.

c. When analyzed by content type, results were mixed, with a few tests showing significance. Trends favored experimental groups.

d. "H_2: Subjects who have not had instruction will experience more difficulty with fallacious arguments than with valid arguments." H_2 was supported by pretest results.

e. Order of difficulty of forms is: modus ponens, contrapositive, converse, and inverse were roughly equal. Invalid forms were considerably more difficult.

f. Content difficulty varied with argument form.

g. No relationship between cognitive style and scores on conditional reasoning was found.

5. Interpretations

"The striking improvement of the experimental groups on the converse seems to indicate that, with instruction, pupils can grasp the principles involved in recognizing the invalidity of this form of argument." The results "strengthened the experimenter's hunch that students tend to confuse the contrapositive and the inverse and that instruction seems to increase the confusion." Results for different content types ran contrary to expectations. There were "more errors on the removed-from-reality fallacious arguments than on the misleading fallacious arguments. One possible explanation for this is that perhaps the subjects were often able to avoid fallacious reasoning by creating counterexamples for proposed incorrect conclusions. If this was their approach, then the situation in the removed-from-reality arguments would hinder them, since it is impossible to create counterexamples in a completely foreign situation." Content appeared to influence difficulty less than form.

Critical Commentary

This study of logical reasoning of low achievers reports little that has not already been reported elsewhere for "average" ability subjects. It appears to be based on the author's doctoral dissertation, dated 1970, and thus in 1976 its discussion of related literature and results is out of date and restricted. While the questions raised here are important in developing a model of logical reasoning skills, assessments of this type have been frequently reported in recent years. The aims of the study

are buried deeply in the article and perhaps more emphasis should have been placed on questions (1) related to differences in reasoning patterns in "low" and "average" or "above-average" achievers, and (2) involved with cognitive style.

The procedures and design raise several questions, among them the following:

(1) H_1 was answered by doing fifteen different comparisons and it is not at all clear how significance on a portion of these would have been interpreted. The analysis looked at whether "more students" improved rather than the suggested (by H_1) "more improvement."

(2) No rationale is provided for the choice of "percent who improve" and the chi square statistic as the means for data analysis.

(3) "Improvement" is defined to mean "'posttest score minus pretest score' is positive." This seems to be a weak condition on a test of 48 items.

(4) Total N varies from 188 to 193 through the article with no explanation.

The graphical presentation of changes in performance for both content type and argument form were well done and help to make sense out of a lot of data. Why the six-group boy/girl split is maintained, however, is unclear unless the schools were considered in some way to be significantly different. Once the fact of no boy/girl differences became apparent, this could have been dropped in the interests of parsimonious reporting.

Conceptually, the discussion of the four well-defined content categories and the hypothesis regarding a relationship between cognitive style and conditional reasoning ability are to be commended. Empirically, the results for the former were mixed, while for the latter came to nought. Nevertheless, these are important concerns. Furthermore, the attempt to improve logical reasoning through instruction, a task seldom undertaken in a rigorous evaluative setting, is likewise to be commended.

References

Ennis, R. H. & Paulus, D. H. Critical Thinking Readiness in Grades 1-12 (Phase 1: Deductive logic in adolescence). Ithaca, N.Y.: Cornell University, 1965 (ED 003 818).

Hill, S. The study of logical abilities of children. (Doctoral dissertation, Stanford University). Ann Arbor, Mich.: University Microfilms, 1961. Order No. 61-1229.

O'Brien, T.D. & Shapiro, B.J. The development of logical thinking in children. American Educational Research Journal, v5, pp531-542, 1968.

Wilkins, M.C. The effect of changed material on ability to do formal syllogistic reasoning. Archives of Psychology, No. 102, 1928.

Lars C. Jansson
University of Manitoba

LEARNING GAMES AND STUDENT TEAMS: THEIR EFFECTS ON CLASSROOM PROCESS. DeVries, D. L.; Edwards, K. J. American Educational Research Journal, v10, pp307-318, Fall 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Robert Kalin, Florida State University.

1. Purpose

To study some effects upon selected classroom variables (such as relationships between students) of using the learning game "Equations", of using student teams, and of using a game-teams combination.

2. Rationale

Previous research had indicated that some learning games:

(a) help increase learning,

(b) establish an "interpersonal experience...due in part to the interdependent task structure which requires interaction among the players",

(c) increase positive reinforcement frequency when played in a tournament structure.

Other research had indicated that some student-team arrangements:

(a) result in students outperforming those working independently,

(b) set up pressures for students to become academically involved so that the team can benefit,

(c) result in higher achievement with teams competing against each other than with each team striving independently for tokens.

There has also been evidence that a combination of inter-team competition with some kind of learning game can result in still greater achievement.

The present research studied the comparative effects of a game and a quiz, of team reward versus individual reward, and of a combination of the game and the team reward. Instead of using achievement as the criterion, it studied the effects of these arrangements upon such "classroom processes" as student interrelationships and learning environment.

3. Research Design and Procedure

One hundred ten seventh-grade students from four mathematics classes in an urban junior high school were stratified as high, medium, and low achievers, then randomly assigned to one of four treatment groups:

(a) Games task with individual reward (apparently 30 students)

(b) Games task with team reward (apparently 30 students)

(c) Quiz task with individual reward (student number unidentified)

(d) Quiz task with team reward (student number unidentified).

The four-week experiment was scheduled as follows:

(a) Monday -- regular classroom instruction (nature, content, and student-grouping undescribed)

(b) Tuesday -- half period of Game/Quiz task, half period of practice.

(c) Wednesday -- regular classroom instruction

(d) Thursday -- half period of practice, half period of regular instruction

(e) Friday -- full period of Game/Quiz task.

Students in the two game-treatment groups played "Equations" (Allen, 1969), with 3 students of similar mathematical achievement per game for a total of 10 "Equations" games per game-treatment group. To maintain homogeneous game tables, students were moved from one table to another according to wins or losses the previous day. Students in the two Quiz-treatment groups were quizzed on Tuesday and Friday over mathematical content studied the previous day.

Students in the two team-reward classes were assigned to four-member teams to as to obtain intra-team heterogeneity as to race, sex and achievement, yet to obtain inter-team equality. They played the "Equations" game or took Quizzes individually. But team scores as well as individual scores were listed in a twice-weekly newsletter for the team-reward group, whereas only individual scores were listed for the individual-reward groups.

The first of three types of dependent variables was the extent of peer-tutoring occurrence, measured by research-assistant observation of two 20-minute practice periods. Each observer made a series of five-second observations of 15 randomly selected students followed by a ten-second recording period. The second type of dependent variable was the extent and direction of friendship and helping patterns, measured by asking students to name their friends, who helped them, and whom they helped. The third type was the student perception of the learning environment, measured by having students respond to four categories of the previously developed and standardized Learning Environment Inventory: Difficulty, Competition, Satisfaction, and Cohesiveness. A newly developed scale of Mutual Concern was also used.

4. Findings

The researchers analyzed the sociometric choice data, the Learning Environment Inventory data, and the classroom observation data separately, using the Goodman technique for multivariate analysis of qualitative data. Using either 5%, 1%, or .1% significant levels, they reported these statistically significant findings:

(a) The games treatment increased peer tutoring, and resulted in classroom process which students perceived as more satisfying, less difficult, and less competitive.

(b) The team-reward treatment also increased peer tutoring, and resulted in classroom process that students perceived as more competitive while allowing greater mutual concern.

(c) The combined games/team-reward treatment additively increased peer tutoring, and created a significant interaction effect as to the classroom climate variables of competitiveness and cohesiveness.

5. Interpretations

Game Effects: The researchers felt that, as compared with the Quiz treatment, the game effect of decrease in perceived difficulty and increase in perceived satisfaction were the result of the "task interdependence created by the game", in which the students learned and played together. An alternative or possible additional cause was that the increased positive reinforcement afforded by the game increased the chances of a student doing well. This reinforcement factor, they felt, may also have caused the decrease in perceived competitiveness, which originally surprised them.

Team-reward Effects: The researchers interpreted as natural their observation that the increase in peer-tutoring, brought upon by the team-reward as compared to the individual reward procedures, appeared to focus upon the help given the low achievers by higher achievers - the team-reward feedback system encouraged concentration on those team members needing the greatest amount of assistance. The result of the team-reward increasing the student perception of competitiveness at the same time as increasing cohesiveness and mutual concern would normally appear contradictory. But such an effect would appear natural when the competition is among teams.

Combined Game Team-reward Effects: The researchers interpreted the increase in peer-tutoring resulting from the combined treatments as created in part by the novelty of the games task.

In summary, they felt that both the "Equations" game and the team-reward technique were helpful in a complementary way in creating new and beneficial classroom processes. Whereas this game helped students view their classroom tasks more positively, the team-reward system encouraged students to work together with mutual concern. The net effect, the researchers felt, accounts for the greater academic achievement claimed for the combination in an earlier study.

Critical Commentary

This study's results agree comfortably with common instructional experience and prior research. For example, one would expect that a ("proven") game like "Equations" would compare favorably with a Quiz treatment when the comparison is in terms of the extent of peer-tutoring, and of student perceptions of which is more satisfying, difficult, and competitive. Full comprehension of this result is difficult, however, since the researchers did not describe the nature or content of the instruction on which the students were quizzed. Did it, for example have the same mathematical content as the "Equations" game?

It would also be expected that the team-reward system would result in an increase in perceived competitiveness over that created by the individual-reward technique. For one reason, each student in the team-reward treatment saw his/her own score as did the student not on a team, then in addition was influenced by his or her team score.

In spite of such expected outcomes, generalization or classroom application should be approached carefully. For one thing, no information as to the nature of the experimental subjects is given beyond their being in an urban junior high school. Of course, generalization should not go beyond the "Equations" game or one with its characteristics, beyond the particular tournament structure used, or beyond the particular type of team-reward structure. Furthermore, the treatments lasted but four weeks; would the effects observed continue over a longer period of time? Most important, there is likely to have been a considerable Hawthorne effect - the experimental conditions were most complicated as well as quite different from the usual classroom procedures.

Robert Kalin
Florida State University

EJ 109 957 SE 512 615

LOGICAL REASONING ABILITY AND TEACHER VERBAL BEHAVIOR WITHIN THE MATHEMATICS CLASSROOM. Gregory, John W.; Osborne, Alan R. Journal for Research in Mathematics Education, v6 n1, pp26-36, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Thomas J. Cooney, University of Georgia.

1. Purpose

The following hypotheses were investigated:

"1. There is significant positive relationship between a seventh-grade student's conditional reasoning ability and his teacher's use of conditional moves.

2. There is a significant positive relationship between a seventh-grade student's mathematical ability and conditional reasoning ability.

3. There is a significant positive relationship between the combination of the teacher's use of conditional moves and the student's mathematical ability and conditional reasoning ability."

2. Rationale

In reviewing related research the investigators made the following observations:

A. Principles of logic can be taught, but students receiving instruction on principles of logic experience approximately equivalent growth in logical reasoning as compared to those not receiving such instruction.

B. Attempts to identify correlates of logical reasoning ability and the developmental process have yielded only the passive variables of I.Q. and age.

C. Teachers do use conditional reasoning in their classes. In particular, Smith and Meux concluded that mathematics teachers use such language more than teachers of other subjects.

D. Students use the verbal behavior of their teachers as a model for their own use of language.

These findings gave rise to the above three hypotheses.

3. Research Design and Procedure

Twenty teachers were randomly selected from eighty-four teachers of regular seventh-grade mathematics in Columbus, Ohio. Five sessions of one class of each teacher were audiotaped and transcribed. The frequency of

teacher conditional moves (based on Smith and Meux's definition of conditional move) were determined by three trained analysts. Intercorrelation coefficients yielded reliability measures of greater than .93 for the three analysts.

Teachers were ranked according to the average frequency of conditional moves per lesson. Ten teachers were then selected, five with highest rank and five with lowest rank. These two groups differed significantly ($p < .005$) on the mean number of conditional moves per lesson. Their means were 31.8 (high) and 11.1 (low). There were 114 subjects in the high classes and 116 subjects in the low classes.

The Cornell Conditional Reasoning Test - Form X was administered the first month of school and again five months later at the end of the first semester to all subjects in the study. The test is designed to measure understanding of various inference schemes, some valid and some not. The test also measures responses to three types of item content:

A. concrete familiar content which the subject has no reason to believe true or false;

B. symbolic content involving letters;

C. suggestive content that is familiar and whose truth status is known to the student (If whales are birds, ...).

"Individual item group reliabilities reported by Ennis et al.(1965) ranged from .27 to .63, with 10 of them .40 or higher. A total test reliability of .85 was determined using scores of subjects in this study."

Students were grouped into three levels of ability (M1, M2, M3). Treatment groups were formed on the basis of the three ability levels and the high (G1) and low (G2) teacher rankings. The number of subjects in each cell is given below:

	M1	M2	M3
G1	17	61	36
G2	11	72	33

The adjusted posttest score, using pretest performance as the covariate, served as the measure of conditional reasoning ability for an analysis of covariance for two-way design and for correlation analysis.

4. Findings

A 2x3 analysis of covariance yielded significant F statistics between the teacher rank groups (G) ($p < .027$) and among the mathematical ability groups (M) ($p < .003$). These results indicated that the relationships expressed in Hypotheses 1 and 2 would be found to exist. No significant interaction (G X M) was found, which led to the acceptance of the null form of Hypothesis 3.

Various significant ($p < .05$) correlation coefficients for independent variables versus adjusted posttest item group performance were identified. "Membership in the high group was found to be related to total test performance ($r = .429$) and performance on the items relative to the principle "p only if q: Not p" ($r = .422$) and suggestive (SU) and negative (NG) content items ($r = .536, .519$, respectively). Mathematical ability was found to be related to total test performance ($r = .473$) and performance on the three fallacy types of denying the antecedent ($r = .450$), asserting the consequent ($r = .646$), and asserting the converse ($r = .521$)." These findings were taken as support for Hypotheses 1 and 2.

5. Interpretations

The significant correlations with individual item groups indicated that the relationships between the two variables and student reasoning ability was differential in nature.

The following limitations were noted: (a) nonrandom assignment of students to treatments; (b) no control for reading abilities, intelligence, or other possible potent variables between treatment groups; and (c) the mathematical content of the lessons was not constant across all teachers. Subsequent analysis indicated that there was no loading in favor of the high group on mathematical ability or on reading ability. There was a significant correlation ($r = .671$, $p < .01$) between reading and conditional reading ability.

Critical Commentary

The line of research taken in this study is particularly cogent in trying to link teacher behavior with certain student outcomes. The investigators correctly emphasized that their correlational design does not allow one to conclude causal relationships. As they pointed out, studies designed to establish causation are needed. The present study provides justification for such research.

There are some concerns associated with studies like the present one which involves observations of teacher behavior. One concern relates to the stability of the behavior being observed. Often variables defined by counts are not highly stable and may vary considerably from context to context. This concern is somewhat magnified in this study since content was not constant for all of the teachers. It appears that the analysis of the teachers' behavior occurred between pre- and posttesting. In any event the question arises as to how representative the observed teachers' behavior was in relation to their behavior over the entire semester (pretest to posttest).

The questions below reflect those identified by the investigators and offer suggestions for future research in this area.

1. If a teacher's use of the language of logic does promote students' understanding of logic, is the students' behavior largely imitative, or does it extend beyond that of imitation?

2. Is the occurrence of language of logic primarily dependent on the teacher or on the mathematical content or on other factors?

3. Is the variable of a teacher's use of the language of logic best described in quantitative terms (e.g., counts) or in qualitative terms defined by the contexts in which the language occurs?

Thomas J. Cooney
University of Georgia

EJ 068 474 SE 507 555

INDIVIDUALIZED INSTRUCTION IN PROBLEM SOLVING IN ELEMENTARY SCHOOL MATHEMATICS. Jerman, Max. Journal for Research in Mathematics Education, v4 n1, pp6-19, January 1973.

Expanded Abstract and Analysis Prepared Expecially for I.M.E. by Gerald A.Goldin, Graduate School of Education, University of Pennsylvania.

1. Purpose

To compare the effectiveness with respect to problem solving performance in mathematics of two individualized programs for the teaching of problem solving skills: (a) the Productive Thinking Program, Series One: General Problem Solving, by Covington, Crutchfield, and Davies, and (b) the Modified Wanted-Given Program, by Jerman.

2. Rationale

It has been conjectured (Crutchfield, 1966) that children may possess general problem solving skills which, when encouraged to develop, are available for transfer to a variety of specific situations and contexts. But in actual schooling these are often neglected, under-utilized, and permitted to atrophy. The Productive Thinking Program was designed to develop existing general problem solving skills in children - how to identify a problem properly, how to look for clues everywhere, how to select relevant information, how not to be afraid of making mistakes, etc. Its context is non-mathematical: children with whom the student can identify solve problems as amateur detectives. The storyline characters are a brother, a sister, and their uncle.

By contrast, the Modified Wanted-Given Program was designed to teach arithmetic problem solving using a mathematical context. A straightforward decision algorithm employs "wanted-given" procedures: the student learns to distinguish between sum-type problems (add or subtract) and product-type problems (multiply or divide); in each case two operations are possible depending on what is wanted and what is given. More complicated problems are solved by nesting such sequences. Storyline characters, a man and a boy, solve arithmetic problems in a variety of situations.

Both programs consist of a self-instructional sequence of sixteen booklets.

It is asked whether mathematics problem solving skills, as measured by a variety of mathematics tests, are more effectively taught by means of the Producting Thinking Program (teaching general problem solving and relying upon transfer), or by means of the Modified Wanted-Given Program (teaching a specific problem-solving algorithm in a mathematical context). Both sequences are to be compared with regular class programs in which problem solving is not taught.

It is further hypothesized that no significant sex differences in achievement would appear within the treatment groups. This is expected due to the programmed, individualized nature of the materials.

3. Research Design and Procedure

Participating in the study were 287 students in eight fifth-grade classes from four public schools in Santa Clara County, California. This is a predominantly white, middle-class socioeconomic area. Data for 26 students classified by their schools as either "gifted" or "educationally handicapped" were disregarded; thus the sample population excluded the extremes of ability. The author employed a 3x2 factorial design, comprising the three treatments (two experimental programs and the regular class program) and the two sexes.

The pretest was administered the day before the treatment period began, and consisted of four scales: Working with Numbers, Arithmetic Reasoning, and Hidden Figures (all NLSMA tests), and Figure Classification (adapted from an ETS test).

The treatment period lasted 16 days, about 40 minutes per day. Students in six of the classes were randomly assigned to one of the two experimental treatments. Booklets were handed out and collected daily by aides in the experimental classes; the classroom teachers did not participate in the administration of these treatments. Teachers in the other two (control group) classes agreed not to give instruction in problem solving during the study.

The posttest was administered the day after the last day of treatment, and consisted of three scales: Working with Numbers and Five Dots (NLSMA tests), and Word Problems (a five-item arithmetic world problem test developed especially for the study). A follow-up test was administered seven weeks later, consisting of three scales: Letter Puzzles I and Directions (NLSMA tests), and a parallel form of Word Problems.

Analyses of covariance were performed using the four pretest scales as covariates with the posttest and follow-up test scales. Later, the two Word Problem tests were re-scored with respect to the use of correct procedures rather than the attainment of correct answers, and new analyses of covariance were performed.

4. Findings

There were no differences significant at the .05 level due to choice of experimental procedure on any of the posttests or follow-up tests. There were no significant differences on any of the posttests or follow-up tests between the experimental groups and the control group. Only the Five Dots posttest had a significant sex effect ($p < .05$) for boys.

When the Word Problems scales were re-scored with respect to the use of correct procedures, there were significant differences on the posttest in favor of the experimental groups over the control ($p < .001$), and in favor of the Modified Wanted-Given Program over the Productive Thinking Program ($p < .005$). On the follow-up test there were still significant differences in favor of the experimental groups over the control ($p < .001$), but the difference in favor of the Modified Wanted-Given Program was not significant. However there was a treatment/sex interaction ($p < .10$) in favor of boys in the Modified Wanted-Given Program.

5. Interpretations

Large numbers of subjects apparently used correct procedures on the word problems, but arrived at wrong answers due to computational errors. Since significant differences did appear on the Word Problems scales when they were re-scored, it is suggested that computational ability may be considerably independent of procedural ability. Further research is suggested to measure the use of correct problem solving strategies separately from computational ability.

Systematic instruction using either treatment was significantly more effective than no instruction, for the use of correct procedures. Thus there was indeed some transfer of training in general problem solving skills to the mathematics word problems. But at least in the short run there was still better performance by children trained in a specific algorithm in a mathematical context, and therefore this approach is deemed more effective.

Critical Commentary

In the case of the Productive Thinking Program, one might attribute the absence of transfer to the high level of generality or even vagueness of the skills being taught, as well as their non-mathematical context. In the case of the Modified Wanted-Given program, one might cite a diametrically opposite reason for lack of transfer, namely the great specificity of the problem-solving algorithm. Consequently it is most interesting that while neither experimental treatment had a significant effect on the posttest or follow-up test scales, both had a significant effect with respect to the use of correct procedures. It is, of course, not surprising that the Modified Wanted-Given Program had the greater effect, since the test problems apparently employed the algorithm taught. Since the word problems were re-scored after the absence of other effects had been ascertained, we would ask whether they were re-scored independently by more than one person, and whether the scorers were "blind" with regard to the treatment administered each subject.

The two treatments seem to represent extremes in level of generality. Future research ought to address the intermediate-level alternative of teaching problem-solving heuristics as described by George Polya and others. Transfer of training in mathematical heuristics ought to be compared with both training in general problem solving and training in specific algorithms.

It is difficult to accept the author's denial of significant sex/treatment interactions. While on any particular scale the sex differences may not be statistically significant, a distinct pattern emerges when the raw score means of all the separate scales are regarded simultaneously. Girls in all three treatment groups scored higher on either 3 out of 4 or all 4 pretest scales. Girls in the Productive Thinking group scored higher on 3 out of the 6 posttest and follow-up test scales, and girls in the control group scored higher than boys on 4 out of 6. However, in the Modified Wanted-Given group, the boys scored higher on all 6 posttest and follow-up test scales, a dramatic reversal of the pretest results. And as already noted, a sex/treatment interaction ($p < .10$)

was reported favoring boys in the Modified Wanted-Given group on the follow-up test in problem-solving procedures.

While the Productive Thinking Program provided both a boy and a girl as role models, it appears that the Modified Wanted-Given Program included only male role models, a man and a boy. This may have unnecessarily weakened the impact of the program on female students, and account for the apparent sex differences in performance.

Gerald A. Goldin
Graduate School of Education
University of Pennsylvania

ED 111 681 SE 019 680

LEARNING OF SELECTED PARTS OF A BOOLEAN ALGEBRA BY YOUNG CHILDREN. Johnson, David C. In Steffe, L. P. (ed.) Research on Mathematical Thinking of Young Children: Six Empirical Studies. Reston, Virginia: NCTM, 1975

Expanded Abstract and Analysis Prepared Especially for I.M.E. by John R. Kolb, North Carolina State University, Raleigh.

1. Purpose

To determine if the ability of young children to (a) form classes and (b) comprehend selected equivalence and order relations is improved by specific training, and to investigate whether transfer occurs to other class-related activities and the transitive property of the selected equivalence and order relations.

2. Rationale

Piaget attributes the growth of intellectual capacity to a developmental process that evolves through four main stages. These stages are: (1) sensory-motor stage, (2) pre-operational representation stage, (3) concrete operations stage, and (4) formal operations stage. At each stage, the child is capable of performing tasks that could not be performed at an earlier stage when the requisite cognitive structure was absent.

When asked to classify, children younger than five usually form "figural collections;" but by age seven, children can sort objects, form unions of classes, and multiply (cross classify) classes. However, genuine operatory classification does not exist until age eight, when children can solve the class inclusion problem (i.e., identify proper subsets and supersets of a given set).

Instruction on forming classes would not be assimilated by children early in the pre-operational stage, according to Piaget, since they do not have the requisite cognitive structure. Likewise, instruction on these tasks would be unnecessary for children who are well into the concrete operational stage since they already can perform these tasks. Therefore, Johnson's investigation deals with improving through training the classifying ability of children who are in transition from one stage to the next. Moreover, he seeks to determine if through transfer the training produces improvement in certain abilities associated with operatory classification and the presence of logical mathematical thought.

3. Research Design and Procedure

The investigator administered the Otis Lennon Ability Test to 196 children in four kindergarten and four first-grade classes in the Athens, Georgia area. The children were classified into four groups using two levels of I.Q. (80-100) and 105-125) and two levels of chronological age

(64-76) and 77-89). These groups were sampled to produce 20 children in each of the four categories. Within each of the four categories, the subjects were randomly assigned to experimental and control groups, producing 10 subjects in each cell.

The experimental groups were provided training in 17 instructional sessions of about 20 minutes each. Through manipulation of physical objects, the children formed (1) classes, (2) intersections and union of classes, (3) the complement of a class, and (4) relations between classes and class elements. The basic connectives used in the learning material to form classes were conjunction, disjunction, and negation. The mathematical relations were the "pre-number" relations of "more than", "fewer than", and "as many as", and the relations "same shape as" and "same color as."

The five basic posttests and what they are designed to measure are:

1. Connective Achievement Test (CA). The ability to interpret "or", "and", and "not" in directions for forming classes from a set of physical objects.

2. Relation Achievement Test (RA). The understanding of the five relations taught in the instructional unit.

3. Multiplication of Classes and Relations Test (MU). The ability to classify, using two or more criteria at once.

4. Class Inclusion Test (CI). The ability to solve class inclusion problems.

5. Transitivity Test (TR). The ability to use the transitive property for each of the relations tested in the RA test.

Test CA and MU were further divided into three subtests each. The first subtest of MU, the three subtests of CA and RA comprised five achievement measures and CI, TR, and the remaining two subtests of MU were designed as transfer measures.

4. Findings

The results of the six MANOVA's and the 27 ANOVA's can be summarized as follows:

1. The differences between the experimental and control groups were significant (.01 or .05) on all achievement and transfer measures except the CI test.

2. The differences between the higher and lower I.Q. groups were significant (.01 or .05) on all achievement and all transfer measures except the third subtest of MU.

3. The first graders performed better, but not significantly better, than the kindergarteners on all achievement and transfer tests.

5. Interpretations

The superiority of the experimental groups on all achievement measures was interpreted to mean that the training produced substantial improvement in forming classes and using "pre-number" relations. Since there were no significant differences in achievement between the two ages on any achievement measure, it was highly recommended that the grade placement of instruction similar to that in this study begin at the kindergarten level.

Although the experimental group performed significantly better than the control group on three of the transfer tests, Johnson points out that the improvement in at least two of these tests may not be a genuine effect of the training but can be explained by other plausible hypotheses. On the crucial CI transfer test, the treatment produced no significant difference and hence no evidence of improvement. Even for the factor of intelligence, which was significant on the CI test, the scores of both higher and lower I.Q. groups were only slightly above the scores expected by guessing and thus operatory classification was not achieved. Johnson concluded that the instruction produced substantial improvement on physical knowledge as measured by the achievement tests, but produced little improvement in logical mathematical knowledge required in the transfer measures, particularly the class inclusion test.

Critical Commentary

The investigator should be congratulated on a carefully devised and well-conducted study. Good planning and attention to every detail are evident at each step. The careful attention to instrumentation and the reporting of test statistics is only one indicator of the overall excellence of the conduct of this study.

Particularly commendable is the discussion section wherein Johnson acknowledges that the significant differences between the experimental and control groups on several of the crucial transfer measures were likely not due to the training but to other factors he observed in the responses of the subjects on these tests.

The intent of this study and other training studies is to determine if instruction will enable a child to successfully perform certain tasks at an earlier age than would be the case without directed experiences. From a research point of view, it is of interest to know exactly what kind of performance can be hastened by training; but from a practical point of view, why expend great instructional effort in instilling a capability that will evolve naturally at a later stage? Moreover, might not the intervention of training at an earlier age have ill effects on cognitive development? Should future training studies concentrate on determining if there are ill effects from such training? Finally, there is a need for more studies like those reported by Bryant (1974) that investigate more precisely why children cannot perform certain tasks such as conservation and transitivity at a particular age.

Bryant, Peter. Perception and Understanding in Young Children. New York: Basic Books, Inc., 1974.

John R. Kolb
North Carolina State University

COMPREHENSION OF NEGATION AND QUANTIFICATION. Just, M.A.; Carpenter, P. A. Journal of Verbal Learning and Verbal Behavior, v10, pp244-253, 1971.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Robert B. Kane, Purdue University.

1. Purpose

Three experiments were designed to study how people process different kinds of negative sentences.

2. Rationale

Six studies are cited which suggest that explicit negative sentences with the word not in them are processed differently from affirmative sentences. This conslusion was based on significantly longer reaction times to such negative sentences than to affirmative sentences. This finding held for the task of verifying as well as completing sentences. The authors note that negation is found in more forms than just the explicit not morpheme. For example, few, scarcely any, and seldom negate implicitly. Moreover, not, none, and no negate a sentence about a complete set, while the implicit negatives like few negate a sentence about a proper subset. They conjectured that processing differences may exist between explicit full-set negation, such as none, and implicit subset negation, such as few. They also conjectured that sentences which have the same referent as few, but which are not syntactically negative, may be processed differently. For example, sentences with a minority refer to the same denotata as those with few, but sentences with a minority are not syntactically negative.

3. Research Design and Procedure

These conjectures were tested by examining how people decode sentences and pictures in verification tasks and represent them internally. Tasks were designed which called for a comparison of information given by a sentence and a corresponding picture. The label true or false (relative to the picture) was to be assigned to each sentence. In Experiment I the S was timed while he read a sentence, looked at a picture, and decided whether the sentence was true or false. Half of the sentences used are listed below; the other half were identical except the word black was substituted for the word red.

Syntactic Category	Affirmatives	Negatives
Syntactic negatives with a negative particle (SYN-NOT)	The dots are red.	The dots aren't red.
	All of the dots are red.	None of the dots are red.
	Three are red dots.	There are no red dots.

Syntactic negatives without a negative particle (SYN)	Many of the dots are red.	Few of the dots are red.
	Most of the dots are red.	Scarcely any of the dots are red.
	Lots of the dots are red.	Hardly any of the dots are red.
Semantic but not syntactic negatives (SEM)	A majority of the dots are red.	A minority of the dots are red.
	About 14 of the 16 dots are red.	About 2 of the 16 dots are red.
	A large proportion of the dots are red.	A small proportion of the dots are red.

Each sentence was coupled with a 4x4 array of dots. For the SYN-NOT category, all 16 dots were the same color. For the other categories the arrays had 14 red and 2 black dots or the reverse. In total there were 72 different sentence-picture pairs. Each pair was viewed by the S in a Cambridge Tachistoscope. The S was timed from stimulus onset to decision about truth or falsity. Ss were 20 English-speaking undergraduates.

The interpretation from Experiment I (see "Findings") was tested in Experiment II by replicating the conditions of Experiment I except to present the dot array prior to presenting the stimulus sentence. Eleven Ss from the same pool were used; none had participated in Experiment I.

In Experiment III, the conditions of Experiment II were replicated except for the introduction of two instructions which told Ss how to represent the picture while they awaited sentence presentation. Only three Ss, two of whom were the investigators themselves, were used.

4. Findings

Experiment I: In general, reaction times (RT) to negative sentences were greater than to affirmative sentences, false sentences had greater RT than true sentences, and RT increased from SYN-NOT to SYN to SEM sentences. Among SYN-NOT and SYN sentences true-affirmatives had shorter RT than false-affirmatives, while true-negatives had longer RT than false negatives. Such interactions were absent for the SEM sentences. This absence for SEM sentences was interpreted as evidence that the S coded the color of the larger subset after reading a SYN negative and the color of the smaller subset after a SEM negative. This difference between SYN and SEM was interpreted as an effect of the two categories on how Ss coded the accompanying picture.

In Experiment II, RT patterns for SYN and SEM were similar. The interaction found in Experiment I was present only (and predictably) for SYN-NOT. Moreover, whatever strategy Ss employed in Experiment I, they were not differentiating between SYN and SEM when they coded the picture prior to seeing the sentence.

Since the data generated by the limited sample in Experiment III seem so questionable, no further reference to Experiment III seems warranted.

5. Interpretations

Both full and subset syntactic negatives focus on the larger subset or entire set while SEM negatives focus on the smaller subset. It appears that all syntactic negatives may have the same basic internal representation as their corresponding affirmatives together with a negation operator, while SEM negatives do not focus on the larger set and are unpreferred only with respect to quantity. Perhaps most generally these experiments confirm earlier findings that negatives require more RT than affirmatives. Even with reading time controlled, syntactic negatives take longer to verify than affirmatives. There is a clear preference for affirmative over negative sentences. Affirmatives are psychologically (and linguistically) less complex than negatives. The data here suggest that even quantifiers referring to small quantities are less preferred than those referring to larger quantities.

Critical Commentary

For the researcher or instructional materials generator in mathematics education the subtleties of the psycholinguistics around which these experiments revolve may not be of primary import. What is clearly important is that from the point of view of comprehension and reader tolerance of instructional prose, there are crucial choices available. Not only should we minimize negative sentences overall, but we might make informed choices among the form of negation when negation is needed. Finally, an important point to retain is the utility of monitoring psycholinguistic inquiry as a continuing data source with which to refine our understanding of prose learning in general and, more particularly, in the language of mathematics.

Robert B. Kane
Purdue University

EJ 075 062 PS 502 529
CONTRADICTION, SURPRISE, AND COGNITIVE CHANGE: THE EFFECTS OF DISCONFIRMATION OF BELIEF ON CONSERVERS AND NONCONSERVERS. Miller, Scott A. Journal of Experimental Child Psychology, v15 n1, pp47-62, February 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by J. Dan Knifong, University of Maryland.

1. Purpose

a. To determine whether conservers change to nonconservers (and vice versa) when confronted with belief-conflicting demonstrations.

b. To determine whether the above conflicts are marked by a reaction of surprise.

2. Rationale

The existence and stage development of invariant mental structures have been suspect since their inception by Piaget. Conservation of weight is but one of many phenomena based on the study of such structures. Assuming that conservation of weight acquisition is a simple matter of learning a scientific fact, several Americans have collected data which seem to contradict the relative stability and hence existence of Piaget's stages. Typically, the predictions of nonconservers (or conservers) are contradicted by placing altered clay shapes on the balance and noticing that the beam remains level (or goes down on one side through a trick). The subject is then retested to see if his predictions are changed.

One might expect such contradictions to prompt disequilibration (which is a hypothesized motivator of cognitive change), and it has been assumed that disequilibration might express itself as a surprise reaction. If such is the case, surprise reactions could solve the problem of discriminating true conservers from pseudo-conservers. Although Piaget has never considered this a problem, it is of concern to others who avoid free-form interview data and seek specific behavioral indicators.

3. Research Design and Procedure

One hundred thirty second, third, and fifth graders were tested for conservation of weight on four items in Phase I of a three-phase interview. For three of the items the interviewer demonstrated that two clay balls were equal in weight. He then re-formed one and asked the conservation question. For the fourth item the interviewer showed a clay ball to weigh more than a clay snake (with the aid of a hidden electromagnet). He then rolled the ball into a similar snake and this time asked a conservation of inequality question.

At this point, 55 subjects who were not consistent (conservers or nonconservers) and 6 nonconserving fifth graders were dropped from the

study. The remaining 69 subjects continued Phases II and III of the interview as either nonconservers or young conservers (second and third graders), or old conservers (fifth graders only).

Phase II consisted of a manipulated weighing of the three pairs of clay objects from Phase I and recording of (a) facial expressions, (b) reaction times, and (c) verbal comments. The electromagnet was used consistently to contradict the subjects' predictions from Phase I. Thus, nonconservers were shown true weighings and the conservers were shown (falsely) the ball as being heavier. Phase III consisted of (a) retesting the subjects' predictions for Phase I items (to measure change), as well as (b) asking for other types of weight predictions, followed by (c) a solicitation of the subjects' view of what had happened. The subjects' explanations during the entire interview were recorded and categorized.

4. Findings

a. There were few differences between young and old conservers.

b. About 49% of the nonconservers and 66% of the conservers (both young and old) changed their predictions following the contradictory weighings.

c. Resistance to the unexpected weighings were more frequent among the conservers than the nonconservers; 22 out of 41 conservers denied the validity of the false weighings.

d. About 86% of the conservers offered at least one explanation of the trick balance which indicated resistance to the nonconserving outcome.

e. An abridged summary of Table 3:

	Proportion of Facial Responses During the Contradictory Weighings					
	Neutral	Serious	Thoughtful	Curious	Amused	Surprised
First Contradictory Weighing	.31	.03	.13	.32	.47	.09

(Some gave varied responses and are recorded in more than one category, except for those who were neutral.)

f. The reaction times to contradictory weighings were roughly double the baseline reaction times.

g. The last test item in Phase I, showing the snake to weigh less than the ball, was missed much more frequently (93%) than the first three (28%) by transitional subjects.

h. A chi square test showed significantly more male conservers than female.

5. Interpretations

The findings are interpreted generally as contradicting Piaget's theory. (a) Piaget's theory would seem to predict a frequent surprise reaction. (b) The experience seemed to teach at least some of the nonconservers to be genuine conservers, although it is suggested that the subject must be ready. (c) Finally, a number of conservers seemed to have given up their belief and returned to being nonconservers, although the short age span might not have included subjects who had "mastered conservation." Further research might include older subjects, allow for a delayed posttest, and/or pursue other Piagetian tasks.

Critical Commentary

Miller has tried to converge Piagetian structuralism and American behaviorism into one study. Unfortunately, he uses altered Piagetian interpretations without explicating the difference. For this reason his negative interpretations only convince non-Piagetians; e.g., if conservation of weight is a concept (in the American tradition) then it might disappear or be forgotten; it might be tested on a pre- and posttest; it might be the logical consequence of some other concept; and it might be mastered in school as are the concepts of addition and the properties of free-falling bodies. But conservation of weight is a Piagetian invention and most of the above locutions violate its parentage. Conservation of weight is not a concept to be mastered like long division; a test may give false clues to its existence; and it is not logically necessary in either the Piagetian or American sense of this term.

A Piagetian might find much cheer in the findings, if not in the interpretations. Specifically:

a. Many conservers showed resistance to the trick balance. Others may have simply discounted this balance as "goofy," altered their predictions to suit the situation, and still retained their position as true conservers.

b. The finding of slow reaction times for the unexpected weighings might serve as the long-sought discriminator for disequilibration and hence true conservation.

c. Similarly, the frequency of expressed curiosity and amusement, as well as surprise (Table 3), might be viewed as combined evidence of frequent disequilibration.

d. Most (94%) of the explanations for nonconservation were either perceptual or blank.

e. Finally, it is difficult to accept the 18 fifth graders as "old conservers" since as many as 6 fifth graders were dropped from the study for being consistent nonconservers. (The data do not report the number of fifth graders in transition, but presumably there were several.)

Until American studies address themselves to Piaget's structuralism, instead of considering his theory as a variant of behaviorism, they will be of less benefit than they might in helping us understand children's thinking.

J. Dan Knifong
University of Maryland

EJ 088 418 SE 509 915

THE INTERACTION OF THREE LEVELS OF APTITUDE DETERMINED BY A TEACH-TEST PROCEDURE WITH TWO TREATMENTS RELATED TO AREA. Montgomery, Mary E. Journal for Research in Mathematics Education, v4 n4, pp271-278, November 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Aaron D. Buchanan, SWRL Educational Research and Development, Los Alamitos, California.

1. Purpose

A. To examine the interaction of two treatments on learning to measure area with three levels of aptitude for learning about measurement units.

B. To study the general effects of experience with measurement units on pupils' subsequent performance in comparing measures where units may not be the same.

2. Rationale

This study grew out of development of curriculum materials apparently in conjunction with the Developing Mathematical Processes Program at the University of Wisconsin Research and Development Center. Some children appeared to be capable of grasping certain unit concepts, while others were not. If procedures could be developed for identifying pupils at different levels of aptitude for learning about units, then different instructional treatments could be developed to improve learning about measurement in general for pupils across aptitude levels. According to the author, researchers have not devoted much attention to children's understanding of the role of units in relationship to the process of measuring.

3. Research Design and Procedure

The design of the study is closely associated with groundwork in aptitude x treatment interactions articulated by Cronbach and Snow (1969). Three levels of pupils' aptitudes for learning about length were identified using a procedure developed as part of the study. Pupils in each aptitude level were randomly divided into two groups and given one of two instructional treatments for learning about the measurement of area. These treatments differed in the amount of emphasis placed on different types of area units in the learning sequence. After the instruction was completed, pupils were tested on their ability to compare measures of area. In some instances the measures to be compared were based on units that were the same (congruent units); in other instances these measures involved units that were not the same (non-congruent units).

The study involved all second- and third-grade pupils (110 subjects) from a single rural community. The procedure used to identify aptitudes for learning about units began with a pretest (20 items) on understanding

units of length. Next, pupils were given two lessons designed to teach them about comparing measures of length, where they must take into account the size of the unit as well as the total number of units involved in each measure. Finally, a posttest (20 items) was given, and pupils were stratified into three levels of aptitude. Level I (low aptitude) included 32 pupils who scored below 55% on both the pretest and posttest and did not improve their scores by more than 10% between tests. Level II included 27 pupils who scored below 75% on the pretest, scored 75% or better on the posttest, and improved their scores by 15% or more between tests. Level III consisted of two pupils who scored 75% or more on both tests. Twenty-nine pupils whose scores did not place them in one of these three aptitude levels (e.g., their pretest score was 75% or higher, but their posttest score was below 75%) were dropped from further study, along with 20 pupils who did not complete the three-day teach-test procedure, and the two pupils in Level III.

The instructional treatments for learning about area were developed as part of the study and extended over nine days. The major purpose of the treatment was for pupils to compare and order measures of area. In Treatment N, 26 pupils (13 each from Levels I and II) were taught how to measure and compare areas which were always covered with congruent units. In Treatment U, 28 pupils (16 from Level I and 12 from Level II) were taught how to measure and compare areas which were covered by units that were not congruent. In comparing two measures, pupils in Treatment U had to consider both the number of units needed to cover a region and the size of the unit. Treatment lessons were taught to all pupils by the same teacher. Four pupils apparently did not complete the treatment sequence.

An achievement test developed for the study was given as soon as treatments were completed, and it was given again, as a retention measure, after four weeks. The test consisted of 30 items requiring pupils to provide a measure of area for a given region, and to compare measures for the areas of two different regions. Some of the comparison items (8) dealt with regions covered by units that were the same (congruent), but most (14) dealt with regions covered by non-congruent units.

A transfer test was also developed for the study and administered at the same time as the achievement test. It consisted of 25 items ostensibly involving "...more complicated questions about area, ...and the attributes of length, capacity, and numerousness."

4. Findings

Treatment main effects and interactions were determined using a 2x2 analysis of variance format for analyzing data. Pupils in Treatment U (where emphasis was on units) performed significantly better on the achievement test given immediately after the treatment, and they performed significantly better four weeks later when the test was given again as a measure of retention. Otherwise, there were no other significant differences on posttreatment measures except for differences attributable to aptitude alone. No significant interaction was found between treatment effects and aptitude.

5. Interpretations

Regardless of aptitude, Treatment U, involving work with units of measurement, was clearly superior to the more simplified instructional sequence for teaching the skills assessed on the achievement test and for helping pupils to retain these skills. In addition, pupils in Treatment U consistently referred to the unit when giving measures for regions, but pupils in Treatment N did not. Also, problems where regions were not covered exactly by units were more difficult than problems where "coverage" was exact, and comparisons with non-congruent units were more difficult than comparisons with congruent units for all pupils.

Critical Commentary

From a technical perspective, this study seems to have satisfied design requirements so that it is statistically meaningful. The attrition of subjects between the teach-test procedure for identifying aptitude levels and posttreatment achievement testing is not completely explained, but it doesn't really seem likely that it would have affected treatment main effects.

As far as aptitude x treatment interactions are concerned, the results of this study only reinforce how elusive the ATI phenomenon is to demonstrate in a research context that is meaningful to classroom learning. To understand pupils' needs well enough to provide appropriate instructional treatments is one thing, but to understand individual differences well enough to engineer classic interactions is quite another. In fact, the whole business of finding a treatment that favorably affects low aptitude performances without similarly affecting high aptitude performances is counter-intuitive, unless there is a ceiling effect.

The results of the study do reinforce two things that we seem to know about learning, but occasionally overlook. The first is that pupils tend to do better when they have had direct practice with a task. From the description provided of the post-treatment achievement test, about half of the items (14 out of 30) involved non-congruent units and were therefore unlike any of the practice received by pupils in Treatment N. On the other hand, they seem to be quite similar to the practice provided pupils in Treatment U. It isn't clear whether pupils in Treatment U had any practice with congruent units, or not. If not, then results suggest that it is easier for pupils to generalize from more difficult (and inclusive) tasks to simpler ones, than it is the other way around: this is the second point we often lose sight of. Transfer of learning from one task setting to another is often much too delicate a phenomenon to survive in robust instructional treatments. In classroom settings, positive transfer seldom works the way we would like. In this study, items on the achievement test surely required a good deal of transfer for pupils in Treatment N, and it didn't happen. When pupils from both treatments were given the test developed specifically to assess transfer, neither group did very well.

Cronbach, L. J. and Snow, R. E. Individual Differences in Learning Ability as a Function of Instructional Variables. (Final Report). Stanford, California: Stanford University, March 1969.

Aaron D. Buchanan
SWRL Educational Research and Development
Los Alamitos, California

EJ 080 319 SE 508 969

TEACHER ATTITUDE AS RELATED TO STUDENT ATTITUDE AND ACHIEVEMENT IN ELEMENTARY SCHOOL MATHEMATICS. Phillips, Robert B., Jr. School Science and Mathematics, v73 n6, pp501-507, June 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Richard Hunkler, Slippery Rock State College.

1. Purpose

To determine if a relationship exists between the arithmetic attitude and achievement of seventh-grade students and the arithmetic attitude of the (a) students' most-recent teachers, (b) students' teachers for two of the past three years, and (c) students' teachers for all three of the past three years.

2. Rationale

The opinion of many mathematics educators is that there is a relationship between teachers' mathematics attitudes and their students' attitudes and achievement in mathematics (DeVault, 1967; Fitzgerald, 1967; Houston, 1967; and Young, 1967). A study which considers the possibility that a student's arithmetic attitude reflects the attitude of the teachers he has had over an extended period of time, rather than the attitude of one teacher whom he has had for a few months, would provide evidence which could be used to help support or negate this opinion.

3. Research Design and Procedure

From a population of 884 seventh-grade students in Cambell County, Virginia, a sample was selected of 306 students who had completed the fourth, fifth and sixth grades in this school district. In addition, 59 teachers from the same district who had taught some of the students during the fourth, fifth, and sixth grades made up the teacher sample.

Apparently the independent variables in the study were (a) the arithmetic attitude of the students' most-recent teacher, (b) the type of teacher attitude encountered by the student for exactly two of his past three years, (c) the type of teacher attitude encountered by the student for his past three years, and (d) student intelligence. Dutton's University of California Arithmetic Scale 3 was used to determine the teachers' arithmetic attitudes, and the Lorge-Thorndike Intelligence Test was used to measure students' intelligence.

The dependent variables were students' arithmetic attitude and students' arithmetic achievement. Students' arithmetic attitude was measured using Dutton's Arithmetic Scale, Form C, and students' arithmetic achievement was measured using the Science Research Associates Series.

Two- and three-way factorial analyses of variance were used to analyze the data. All of the factors used in each analysis were not specified.

4. Findings

a. Most-recent-teacher arithmetic attitude was significantly related to students' arithmetic attitude, but not to students' achievement ($p < .05$).

b. Type of teacher arithmetic attitude encountered by the student for exactly two of his past three years was significantly related to students' arithmetic attitude and achievement ($p < .05$).

c. Type of teacher arithmetic attitude encountered by the student for all three of his past three years was significantly related to students' arithmetic attitude and achievement ($p < .01$ and $p < .05$, respectively).

d. There was no significant interaction among the variables of (1) type of teacher arithmetic attitude (for one, two, or three years), (2) student arithmetic attitude, and (3) student intelligence.

5. Interpretations

The author believes that this study lends support to the opinion that a teacher's arithmetic attitude is significantly related to the student's attitude and achievement in arithmetic. Furthermore, he believes it shows that students' attitude and achievement are related in a positive way to the favorable attitudes of two of his past three teachers and to the favorable attitudes of his last three teachers.

Critical Commentary

Designing a study to consider the possibility that students' attitudes result from the attitudes of several teachers to whom the students have been exposed over an extended period of time, rather than from the attitude of a single teacher, is an excellent idea. However, this reader is unable to form his own conclusions about the internal validity and results of this study because of the lack of information. Among the most important omissions occurring in this paper are:

1. Neither the student attitudinal mean scores nor the student achievement mean scores were presented for any of the analyses.

2. Only one factor in each of the two- and three-way factorial ANOVAs was identified (Tables I, II, III, IV, V, and VI).

3. The levels of each identified factor were not spelled out. For example, in the first two-way factorial ANOVA table involving the independent variable, most-recent-teacher arithmetic attitude, one must conjecture from the "Findings" and the degrees of freedom in the ANOVA table that the three levels of teacher attitudes are "Favorable", Neutral", and "Non-favorable". However, the answer to the question, "What teacher arithmetic attitudinal scores were used to operationally define these

categories?" is left to the imagination of the reader. This is also true of the single identifiable factors in the other ANOVA tables (Tables II, III, IV, V, and VI). As to the levels of the unidentified factors, one is not even at liberty to speculate.

4. Since the researcher did not explain or report the results involving all the factors in the analyses, why was a factorial design used? The reason for global use of this design was never stated. One can understand the use of a factorial design to substantiate the findings concerning the interaction of the independent variables, but what about the analyses which did not identify all the factors, making it not only impossible to identify the interaction, but also impossible to understand its purpose?

5. No mention of the need or use of a Multiple Comparison test was made in connection with the ANOVA tables in which the degrees of freedom exceeded one and the F-ratio was significant (Tables I and II).

6. In ANOVA Tables I, II, III and V, the denominator of each of the computed F-ratios was the Within Groups Mean Square; however, in ANOVA Tables IV and VI the denominator of each of the computed F-ratios was the Within Cells Mean Square. Is the reader to assume from the analyses involved in Tables I, II, III, and V that all the factors are fixed, whereas for the analyses involved in Tables IV and VI the factors are random? Why are similar factors in one instance considered fixed and in the next instance considered random?

The idea for the research was a good one. However, this reader must question how effectively the author used acceptable research design and statistical techniques to conduct his research since many important aspects of the study were not reported.

Richard Hunkler
Slippery Rock State College

EJ 068 641 SE 507 557

RETENTION OF PROBABILITY CONCEPTS: A PILOT STUDY INTO THE EFFECTS OF MASTERY LEARNING WITH SIXTH-GRADE STUDENTS. Romberg, Thomas A.; Shepler, Jack. Journal for Research in Mathematics Education, v4 n1, pp26-32, January 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Paul R. Trafton, National College of Education.

1. Purpose

To investigate the retention by sixth-grade students of probability content on which mastery was demonstrated.

2. Rationale

The study is part of the assessment of a unit on probability and statistics. It is an extension of previous research (Shepler, 1970) and analyzes student performance on a retention test.

Studies of retention frequently reveal considerable forgetting occurring soon after initial learning. On the other hand, it has been demonstrated that retention is greater when the degree of initial learning is great. In particular, mastery learning approaches, in which emphasis is placed on having most students learn at a high level, can be viewed as promoting high level of retention.

Since the initial study reported strong results [mean of 66.8 (92.8%) out of 72 items, mastery of 11 of 14 objectives, correct response to 90% of items by 21 of 25 students], the retention test provides an opportunity to examine the hypothesis that the retention level of mastery learning should be high.

3. Research Design and Procedure

Subjects were 25 sixth-grade students who were taught a unit on probability and statistics. A 72-item test measuring 14 behavioral objectives was administered as a pretest, posttest, and retention test. Test items dealt with one-dimensional sample spaces (36 items), two-dimensional sample spaces (19 items), one- and two-dimensional sample spaces (7 items), and ordering two fractions (10 items). Models described or pictured were coins, dice, spinners and boxes containing objects. The retention test was administered 4 weeks following the posttest. In this interval no instruction or practice on probability concepts occurred. Performance on the retention test is analyzed informally and compared with prior performances of the test for individuals and the group.

4. Findings

A. Correlation between initial learning and retention.

The correlation between the posttest and retention test was .78. On the posttest 21 of 25 students responded correctly to 90% of the items (with all students reaching 80%), while on the posttest 17 students reached the 90% level (6 others reached the 80% level). Retention ratios (amount retained ÷ amount learned) varied from 60% to 105% with 21 students having retention ratios in the range .94 to 1.05.

B. Amount of retention by total group.

1. Overall Test. The results for three administrations of the test are:

	Pretest	Posttest	Retention
Mean (72 items)	27.28	66.80	64.40
Mean in terms of percentage	37.9%	92.8%	89.5%
Variance	74.13	11.17	56.95
Retention ratio			.96

2. Objectives: On the posttest subjects attained the 90% level (90% of students demonstrated mastery) for 11 of 14 objectives. On the retention test, performance on 7 of those 11 objectives remained above criteria levels. Retentions ratios for objectives ranged from .43 to 1.09 with ratios $\geq$ .80 for 11 objectives. The two objectives for which initial achievement was low had low retention ratios.

3. Items: Retention ratios were above .90 for 63 of 72 items.

5. Interpretations

The results support the hypothesis that a high level of initial learning (90% mastery by 90% of the students) produces high retention levels. This holds for individuals as well as group performance on specific objectives and items.

The authors caution against generalizing the results. They note that the design did not control for test-treatment interaction, although they suggest that it would be unlikely for the large item gains and high retention ratios to be due to this. They also raise the question of the length of retention periods for measures of forgetting having practical significance, and recommend the use of longer retention periods in future studies.

Reference is made to two researchable problems emerging from the study. The first is the relationship between complexity of objectives and ability level of students. The two objectives on which initial performance and retention performance was low were ones for which such results might have been predicted from a Piagetian framework. Second the high correlation for a test with low variance suggests the hypothesis that "...with students who learn for mastery, questions missed on an exam

represent material which is not known. In that case, the high correlation observed would be understandable."

Critical Commentary

This informal study has much to recommend it. It suggests a procedure for guiding curriculum development and validating materials. The approach to assessing and analyzing learning provides helpful ideas for use in more carefully controlled studies. It provides a small additional piece of evidence and much food for thought on an interesting hypothesis that deals with important aspects of learning and instruction.

Three additional observations can be made about the study:

1. Although data about the instructional unit and the development, validation and reliability of the test are provided in earlier research reports, a summary of such information here would help the reader interpret the findings more adequately.

2. The high pretest scores (40% of subjects attained 50% or greater) should receive greater attention in the discussion since much of the analysis deals with the comparison of learning occurring during instruction and the retention of that learning. The attainment of mastery learning status may have been a function of knowledge brought to the study of the unit. Perhaps the test was too easy, measuring too heavily what was previously known and thus not reflecting adequately what was learned during instruction. This could account for the strong retention test performance. In any event it weakens the investigation of the relationship of the effects on retention of initial learning under instruction at the mastery level. For the four students who performed the poorest on the pretest, their performance on the retention test seems to fit better their pretest performance rather than their posttest performance, even though two of them showed mastery on the posttest. It would seem that future investigations of the central question could profit from controlling more carefully knowledge students bring to the study of new content.

3. The test had 72 items measuring the 14 objectives. Yet three objectives had only one item each (two objectives each were measured by two items). One might question whether this is adequate to determine mastery learning of an objective. Two of the three objectives measured by one item showed low performance on the posttest and retention test. Yet these objectives were the basis of interpretation and discussion regarding the relationship between levels of cognitive development and learning. With such a small information base caution needs to be observed in any interpretation of performance.

Paul R. Trafton
National College of Education

THE CHILD'S CONCEPT OF LIMIT. Taback, S. In Rosskopf, M. F. (ed.), Children's Mathematical Concepts: Six Piagetian Studies in Mathematics Education. New York: Teachers College Press, 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by John G. Harvey, University of Wisconsin-Madison.

1. Purpose

This study investigated children's intuitive understanding of several concepts subordinate to the concept of limit. The subconcepts are (1) the rule of correspondence for the terms of a sequence, (2) the neighborhood of a point, (3) convergence (divergence) of a sequence, and (4) limit point.

2. Rationale

In The Child's Conception of Space (1967), Piaget and Inhelder describe a continuity experiment in which they investigate children's ability to conceptualize the iterative subdivision of one- and two-dimensional geometric figures and the limit of the resulting sequence of objects. In addition, children were asked to insert as many points as possible between two points on a line segment in order to discover if the children think of the line as a collection of points. In this study the investigator acknowledges that this work of Piaget and Inhelder does provide an introduction to children's understanding of the concept of limit, but he feels that it does not give a true representation of that understanding because ". . . (1) he (sic, Piaget and Inhelder) relied solely on the process of subdivision to obtain his results and (2) he relied solely on objects and situations that have a distinct geometrical orientation." This study attempts to extend the initial work of Piaget and Inhelder. The investigator feels that his study is especially relevant because mathematics educators have recommended an earlier introduction of topics in analysis, especially the distinction between rational and irrational numbers and infinite sequences of real numbers.

3. Research Design and Procedure

Eight tasks were designed to implement the study. The tasks, each named for some pertinent physical object, are (1) Magazine Covers Task, (2) Alternating Rabbit Task, (3) Diverging Squares Task, (4) Converging Streets Task, (5) Neighborhood Task, (6) Halfway Rabbit Task, (7) Steps Task, and (8) Probabilistic Rabbit Task. In all but the Neighborhood Task each S was presented with a situation in which a given rule defined a sequence of converging or diverging objects; S had to indicate whether he understood the rule. Next, the S was asked to demonstrate, at both the concrete and abstract level, whether the sequence converged or diverged. Finally, if the sequence converged, the S was asked to identify the limit point. The Neighborhood Task appears to be quite similar to those used by Piaget and Inhelder (1967).

A pilot study was conducted in which each of the tasks were administered to 30 Ss. On the basis of this pilot study it was decided to restrict the Probabilistic Rabbit Task to Ss in the main study who had "clear understanding" of the limit point in the Halfway Rabbit Task.

Five schools participated in the main study. Five children were selected at random from each school at each of three age levels--eight years, ten years, and twelve years. Almost all of the Ss came from middle to upper-middle class homes. The investigator describes them as ". . . verbal, unusually sophisticated, and very much at ease in expressing their opinions." The tasks were individually administered to each of the 75 Ss. A tape recording was made of each interview and evaluated using a predetermined rating scheme. The categories used to rate Ss responses were (1) clear understanding, (2) uncertain understanding, and (3) no understanding. The investigator reports that subject performance was generally not difficult to categorize; an answer was usually completely right or completely wrong. If a Ss responded that he did not know the answer or if he failed to respond, the response was classified as being in Category 3. In order to measure the reliability of the rating scheme a colleague of the investigator was trained to rate the audio tapes; there was a 93.2 percent agreement between the two judges. The number of responses in each category for each task are reported for each of the four subconcepts.

4. Findings

Except for the Probabilistic Rabbit Task, all Ss attempted the first part of every task, but not necessarily the other parts of the tasks. On each of the four limit subconcepts, tables are included which describe the number of Ss classified as being in each category on each task. Although the study was not designed to study differences between girls and boys, the data were retabulated to present these results separately. Aside from this, no further analyses of the data were performed.

In order to condense this data for presentation here, the percentages of Ss who responded on some task are presented by category. Each percentage was calculated by totaling the number of responses rated to be in a given category and dividing that number by the total number of responses on all the tasks related to that subconcept. These calculations resulted in the following:

SUBCONCEPT	CATEGORY 1			CATEGORY 2			CATEGORY 3			Total No. of Responses		
	Age			Age			Age			Age		
	8	10	12	8	10	12	8	10	12	8	10	12
Rule of Correspondence	52%	81%	81%	13%	9%	13%	35%	10%	6%	150	150	150
Convergence												
Concrete Level	64%	72%	80%	8%	5%	2%	28%	23%	18%	117	158	165
Abstract Level	51%	80%	88%	17%	5%	2%	31%	15%	9%	70	98	112
Neighborhood												
Open Interval	12%	56%	64%	48%	24%	32%	40%	20%	4%	25	25	25
Open Circle	18%	38%	44%	27%	33%	28%	55%	29%	28%	22	24	25
Limit Point	33%	43%	50%	18%	12%	9%	49%	44%	41%	72	108	117

5. Interpretations

The investigator concludes that eight-year-old subjects could do little more than follow a simple rule of correspondence. He notes that usually a maximum of 60 percent of them were successful with the rule of any particular task. Next he concludes that the eight-year-old subjects, even when they were able to understand the convergence of a sequence on the concrete level, were unable to do so on the abstract level. Finally, within this age group the investigator comments that 20 percent or less of the children could indicate the boundary of a neighborhood or identify any single limit point.

On the basis of the data the investigator concludes that the 10-year-olds' performance was very similar to that of the 12-year-olds on most of the concepts examined. In strong contrast to the performance of the 8-year-olds those 10- and 12-year-olds who understood the convergence of a sequence on the concrete level were usually able to do so on the abstract level. While 10- and 12-year-olds performed almost identically in describing the boundary of a neighborhood, they performed quite differently in conceptualizing the number of points within it. He states that ". . . only one 10-year-old child could conceive of infinitely many points, where eight 12-year-olds could do so." He also concludes that the older subjects were able to conceptualize a given limit point to a greater degree when the point was not actually visible in the diagram.

Finally, the investigator concludes that all three age groups included in the study performed at extremely low levels compared to the Geneva group.

Critical Commentary

This study appears to have been well conceived, well planned, and executed with care. However there are some questions raised by this study which should be asked and carefully considered. They are:

1. Why did not the investigator first initiate a study which attempted to replicate the continuity experiments of Piaget and Inhelder? Poor performance of U.S. children on those tasks might have led the investigator in directions other than the one reported here, but one could speculate that had it not, this study would have been greatly improved.

2. Why did the investigator focus primarily on sequences and sequential convergence? If one looks ahead in the mathematics curriculum, the more important process would seem to be one which analyzes what is happening to the function f whenever points in its domain are in a neighborhood of a point a. The limiting process applied to sequences is not necessarily the first step in acquiring the limiting process.

3. What opportunity to learn about limits and neighborhood had been afforded the children who participated in the study? While the performance of the children on the tasks indicates they had had little opportunity to learn, this is not documented. And if, as he states, the investigator is interested in whether or not children can acquire these concepts, why did he not generate some learning situations embodying these concepts and add to his investigation by studying the effects of that treatment? The present study does not really help mathematics educators to determine if it is appropriate to attempt to teach these concepts to 8-, 10-, or 12-year-old children; at best it documents that these concepts are not generated spontaneously from the child's experience.

4. Some of the conclusions seem strained. Too often he makes conclusions based on responses from fewer than one-third of his subjects of a given age. While these conjectures are the substance of good research hypotheses, they seem to strain the quality of sound inference.

John G. Harvey
University of Wisconsin-Madison

THE RELATIONSHIP OF AREA CONSERVATION TO AREA MEASUREMENT AS AFFECTED BY SEQUENCE OF PRESENTATION OF PIAGETIAN AREA TASKS TO BOYS AND GIRLS IN GRADES ONE THROUGH THREE. Taloumis, Thalia. Journal for Research in Mathematics Education, v6 n4, pp232-242, November 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Leslie P. Steffe, University of Georgia.

1. Purpose

Taloumis investigated the relationship between area conservation and area measurement. Specific questions posed by Taloumis were: (a) do children who score higher in area measurement also score higher in area conservation, or (b) in accordance with Piaget's theory, would the concept of conservation of area have to be attained to facilitate that of area measurement?

2. Rationale

Taloumis indicated that even though several studies had investigated either area conservation or area measurement, at the time of the study, no research had investigated the relationship of the level of attainment of area conservation and the level of attainment of area measurement.

3. Research Design and Procedure

The sample consisted of 168 middle-class children ranging in age from 6 years 5 months to 9 years 4 months selected from grades 1, 2, and 3 (56 from each grade). Twenty-one classrooms of children from 3 schools were used as a population, 7 at each grade level. Classrooms were selected randomly within schools after which children were selected randomly within classrooms--4 girls and 4 boys from each.

Three area conservation and two area measurement tasks were presented to each child on an individual basis. Two task sequences were used; $C_1C_2C_3M_1M_2$ and $M_1M_2C_1C_2C_3$, where C_1 is conservation task 1 and M_1 is area measurement task 1. One half of the 168 children were administered one sequence and one half the other.

Area conservation task C_1 involved physical transformation of one of two rectangular cards by cutting, separation, rotation, and rearrangement. Task C_2 involved placing "rocks" on "fields of grass" in differing arrangements. Task C_3 involved two "grounds" for flowers being compared as one was successively rearranged.

Measurement task M_1 involved the concept that the areas of two noncongruent regions may be compared by determining the number of unit-squares covering each of the regions. Task M_2 was essentially the same as M_1 except for intratask sequence of unit coverings.

ANOVA was used to analyze the data. The factors were Grade, Sex, and Sequence. Regression analyses were also done for each grade with either Conservation or Measurement as criterion variable and Sex, Sequence, and Conservation or Measurement (depending on which was used as criterion) as predictors.

4. Findings

1. Using Conservation as the dependent variable, Grade and Sequence were statistically significant.

2. Using Measurement as the dependent variable, Grade and Sequence were statistically significant.

3. The mean for the conservation tasks in the case of CM (conservation tasks presented first) was 48.20 and in the case of MC, 51.81.

4. The mean for the measurement tasks in the case of CM was 51.43 and in the case of MC, 48.62.

5. Interpretations

Taloumis included the following points in discussion of the study.

1. If area conservation tasks are administered first, the scores on the area measurement tasks are increased, and vice versa.

2. Piaget's theory . . . that the ability to measure quantities is dependent on acquired concepts of conservation, appears not completely tenable.

3. . . .area conservation performance and area measurement performance are equally predictable if the scores on one of these is known.

4. . . .children's performance in area conservation and area measurement show no differentiation in amount of increase from grade to grade. . . . These results do not agree with Piaget's theory that measurement occurs "spontaneously," . . . .

5. . . .the study lends support to programs that include instruction on area measurement in the primary grades . . .

Critical Commentary

It seems to be popular these days to report studies of children's basic mathematical concepts which appear to be stimulated by some issue raised by Piaget's theory of cognitive development. The present study raises an issue of the psychological relationship between conservation of area and measurement of area. The issue can be clarified by an analogy. Consider the group of Euclidean motions of the plane. A geometrical property is called Euclidean if it is preserved under the group

of Euclidean motions. For example, distance is a Euclidean property, as it is preserved under glides, reflections, rotations, and their compositions. Distance is not a projective property because projectivities exist which do not preserve the distance between two points. In fact, it makes little sense to think about distance in the projective plane. The preservation (or conservation) of a geometrical property, then, is relative to a group of transformations (in mathematics). It makes little sense to talk about preservation of a property apart from a group of transformations.

In cognitive development psychology, the situation as I understand it is quite analogous. Conservation (or the lack of it) is manifestation of presence (or lack) of groupings (logical or infralogical) in the cognition of the individual. Data on conservation of area and measurement of area as measured by Taloumis ought to be statistically related. That it was is not surprising. One does not acquire "conservation" and then acquire "measurement" if Piaget's theory is to be believed.

There are also issues concerning competence and performance that Taloumis does not address. Such distinctions could shed light on why non-uniform jumps did not occur in the data across ages.

In view of literature indicating sex differences in mathematics, it is important to note the nonsignificant results due to Sex. Could it be that differences observed at older age levels are not indicative of a capability to learn mathematics?

Leslie P. Steffe
University of Georgia

STUDY II--AN INVESTIGATION OF THREE DIFFERENT RESPONSE FORMATS FOR A TWO-DIGIT ADDITION AND SUBTRACTION TEST. Wearne, Diana; Whitaker, Donald; Romberg, Thomas A. In Romberg, Thomas A. (ed.), Four Studies on Aspects of Assessing Computational Performance. Technical Report No. 297. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin-Madison, July 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Barbara J. Pence, Stanford University.

1. Purpose

The purpose was to determine if the response format for a 30-item two-digit addition and subtraction test affects the score. The formats investigated are completion, multiple-choice, and shade-the-circle.

2. Rationale

In assessing computational skills, the completion format is the most desirable but time and number of children frequently dictate the use of machine-scorable tests. The most common machine-scorable test format is multiple choice. Disadvantages of the multiple-choice format include guessing, reading problems, confusion produced by the variety of choices, clues given by the choices, and limited use for diagnostic purpose.

Despite the disadvantages, several advantages for machine-scorable tests explain the popularity of the multiple-choice format: the scoring is less expensive, more accurate, and requires less time (Bloom et al., 1971). Thus an investigation of the effect of different test formats including a machine-scorable format which avoids the short comings of the multiple-choice format is needed.

3. Research Design and Procedure

Three response formats were constructed for a 30-item 2-digit addition and subtraction test in which two-thirds of the items involved regrouping. All formats used the same items in the same order. The formats were:

(a) Completion. The child recorded his response below the item.

(b) Multiple choice. Four choices were provided for each item, one of which was "none of these."

(c) Shade the circle. The response format consisted of a three-column grid in which each column contained the digits from 0 through 9. The child shaded the circles in the appropriate column corresponding to the digits in the solution.

Sixty-one second-grade children from one school were randomly assigned to three groups. Each group received a different form of the test; the machine-scorable formats were administered by the authors of the study and the completion format was given by a classroom teacher.

Analysis of the results included computation of means and reliabilities, pair-wise comparisons of the means, and an analysis of variance.

4. Findings

(a) Means and reliabilities. Means varied from 16.90 for multiple-choice to 18.86 for shade-the-circle. Hoyt reliabilities varied from .83 for completion to .94 for shade-the-circle.

(b) Pairwise comparisons of means. Differences between means for multiple-choice vs. completion and shade-the-circle vs. completion were not significant ($p < .6037$ and $p < .5312$ respectively).

(c) ANOVA. The overall F-ratio was not significant ($p < .5005$).

5. Interpretations

For this group of children, the test response format did not influence the score on the computation test. Thus the authors believe that for a test of computational skills, the shade-the-circle format provided a viable alternative to the multiple-choice format.

Critical Commentary

This investigation was well designed and clearly described; a realistic and manageable question was asked and studied. One problem arises concerning the data. The means were recorded for each experimental group; the standard deviations, however, were omitted. Without standard deviations, it is impossible to pursue the validity of the analyses.

Research literature contains many replications of these findings for the comparison of the completion format and multiple-choice format. Studies have indicated that when students are given each format, completion and multiple-choice, testing the same materials, the scores are very similar (Bloom *et al.*, 1971).

The significant contribution of this study is the examination and derived viability of the shade-the-circle format for computation tests given to early elementary school children.

The significance of the study's conclusions is limited. Extension of the investigation to other age populations and cognitive levels would be valuable. Alternative machine-scorable formats which would minimize

the disadvantages of the multiple-choice format for comprehension and analysis items would be especially interesting. Diagnostic test formats would define an additional domain for further research.

Barbara J. Pence
Stanford University

EJ 119 988 SE 514 003

INTERACTION OF SPATIAL VISUALIZATION AND GENERAL REASONING ABILITIES WITH INSTRUCTIONAL TREATMENT IN QUADRATIC INEQUALITIES: A FOLLOW-UP STUDY. Webb, Leland F.; Carry, L. Ray. Journal for Research in Mathematics Education, v6 n3, pp132-140, May 1975.

Expanded Abstract and Analysis Prepared for I.M.E. by Clyde A. Wiles, Indiana University Northwest.

1. Purpose

The purpose of this study was to inquire further into an aptitude-treatment interaction reported by Carry (1968). Carry's study indicated an interaction between the aptitudes of spatial visualization and general reasoning with instructional treatments described as graphical and analytical. However, the reliability of the criterion test was very low, and the direction of the interaction was opposite from that anticipated. The specific purpose of the present study was to replicate the earlier finding, if possible, and at the same time extend and clarify the meaning of the unexpected interaction.

2. Rationale

The theoretical base for this study involved Melton's (1967) model for investigating individual differences in learning. Melton's model includes four major components, each with several subcomponents, that describe the learning process. Briefly, the four major components are: 1) the stimulus-differentiation (the external stimuli together with the internal encoding response of the learner), 2) the response integration (the totality of learner responses to the stimuli), 3) a "hook-up" between the first two components, and 4) an association-mediation component that represents an alternative mediational route which presumably functions until the "hook-up" is formed. The study being reviewed is concerned with the association-mediation component and one subcomponent of the response integration.

It was hypothesized that subjects with relatively high spatial visualization aptitudes have association-mediation components that would be expanded by the verbal analytic treatment, and that subjects with relatively high general reasoning aptitudes have association-mediation components that would be expanded by their graphical treatment. Students experiencing a treatment "matched" with their aptitudes would not be expected to expand their association-mediation components.

Therefore, it was hypothesized that spatial visualization would predict success on the transfer test in the analytical treatment, and that general reasoning would predict success on the transfer test in the graphical treatment.

The appropriate measure of this expansion of association-mediation components was taken to be a transfer test.

3. Research Design and Procedure

The experiment involved 249 second-year algebra students who were randomly assigned to the two treatments. Each treatment involved four periods of unspecified length on presumably consecutive days. The first was spent completing four measures of aptitude, the second and third were spent reading programmed instruction (PI) booklets, and the fourth period was spent working through the transfer test.

The four aptitude tests were obtained from French, Ekstrom and Price (1963) and Guilford (1952), and involved 100 items. Two of the tests were measures of spatial-visualization and two were measures of general reasoning. The criterion test was a 20-item revision of Carry's (1963) test. It is described as containing "new problems that are related to the problems presented in the PI booklets."

The PI booklets dealing with the solving of quadratic inequalities were those "used by Carry (1968) -- revised, rewritten, and extended." One of these, the analytical treatment (A), was "designed to use the multiplicative property of signed numbers," and the other, the graphical treatment (G), was "designed to make maximum use of graphs."

The report does not indicate the number of frames, problems, or length of the PI booklets. Measures relating to student success in completing the booklet were not reported.

The mean scores on the transfer tests were expected to be the same. The appropriate test for an interaction was taken to be an F test for homogeneity of regression in the hyperspace.

4. Findings

The findings were as follows:

(a) A visual comparison of the group means on the criterion tests (9.07 and 8.60 with standard deviations of little more than 3) led to the conclusion that the means were the same for both groups.

(b) The reliability coefficients of the transfer test were greater than those obtained previously by Carry.

(c) No interactions were indicated by the F test. In summary, there was no evidence of an interaction of the instructional treatments with the aptitude variables.

5. Interpretations

The data did not support the findings of the earlier Carry study. In fact, two common items from the Carry criterion test and the criterion test of this study produced inconsistent data.

While acknowledging that the results of this experiment were not promising, the authors indicated continued optimism about ATI studies in genera

Critical Commentary

A strength of this study is the attempt to employ a theoretical model to clarify the instructional factors actually present in the event an ATI is found. Furthermore, in light of the frequent failure of studies to find aptitude interactions at all, the need to replicate when such findings are reported is apparent. Aside from this, a number of issues came to mind during the course of preparing this expanded abstract.

1. It is not clear what transfer means in this study. The concern with reliability is evident, but what can be demonstrated of its validity?

2. Exactly what did the treatments teach the subjects? Was there differential learning of content due to treatments? Did the treatments attempt to teach exactly the same objectives? Whatever the case may be, it seems unreasonable to expect much change in students' abilities to do transfer tasks if they have not shown a mastery of some other related tasks. It seems not too unlikely that the extensive pre-testing would teach as much as the two periods spent reading programmed instruction booklets.

3. It is not clear what implications the study has for the theoretical model. If an interaction similar to that found in the earlier study has occurred, the model would surely have been appealed to for an explanation. Since no interaction was found, do we then conclude that the model is faulty, or perhaps that this application is faulty, or some other alternative? One usually thinks in terms of matching instructional treatments with observed abilities when ATI studies are discussed. The idea of choosing a deliberate mismatch of instruction and aptitude with the objective of compensating for observed ability weaknesses seems novel. Perhaps this idea should be explored in greater depth.

4. But suppose the treatments were clearly distinct and that both were adequate in promoting mastery of some clearly identified objectives. And suppose furthermore, that the criterion test was a valid test of some clearly defined transfer task; what then? It is not clear that we would have much more information. Perhaps we err in supposing that any adequate instructional treatment, even carefully circumscribed programmed ones, can be sufficiently unidimensional so as to match (or mismatch) identified abilities in such a way that other unidentified abilities do not confound our intentions in some way. To the extent that this may be true, populations with severely limited abilities may provide the clearest direction for ATI studies.

69 Clyde A. Wiles
Indiana University Northwest

EJ 090 022 PS 503 012

PIAGET'S SYSTEM OF 16 BINARY OPERATIONS: AN EMPIRICAL INVESTIGATION. Weitz, Lawrence J.; Bynum, Terrell Ward; Thomas, James A.; Steger, Joseph A. The Journal of Genetic Psychology, v123, pp279-284, 1973.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Suzanne K. Damarin and Richard J. Shumway, The Ohio State University.

1. Purpose

To replicate, using a larger group of subjects, results reported by Inhelder and Piaget (1958) concerning the use of 16 binary operations by an individual at the stage of formal operations.

2. Rationale

Inhelder and Piaget maintain that subjects at the stage of formal operations use the full lattice of 16 operations on pairs of statements in the process of solving problems, while subjects at prior stages do not. Their evidence lies mainly in the analysis of one protocol. In earlier work three of the current authors re-analyzed this protocol and identified only six of the operations. Moreover, two professional logicians among the authors were able to construct natural language equivalents for only 10 of the 16 logical forms (operations).

3. Research Design and Procedure

Fifty-seven subjects at three age levels (9-, 12-, and 16-year-olds) were randomly selected from the population of a school system. Each subject was tested individually using the Invisible Magnetism apparatus, which was built in conformity with the descriptions given by Inhelder and Piaget. Subjects were asked to find and verify the reason that a rotating iron needle stopped at a certain position; several irrelevant variables were present in addition to hidden magnets. A system devised for the purpose was used to score protocols for the number of uses of each of the 10 logical operations for which natural language equivalents were identified. Differences in usage frequencies were tested using the chi-square statistic.

4. Findings

Subjects in the three age groups were not found to differ in the number or types of logical operations used. All groups employed the same five operations: conjunction, implication, disjunction, converse implication, and independence of p in relation to q; the first two of these forms accounted for a majority of the logical usages observed. However, older subjects used the 5 forms in more complex and integrated ways than younger subjects did. Eighty percent of the subjects tested immediately guessed the role of magnetism.

5. Interpretations

The authors question the validity of Inhelder and Piaget's interpretations, and conclude that logical growth would be measured more effectively by instruments using only the most common logical forms in situations of graded complexity.

Critical Commentary

1. Although the decision-making procedures for classifying instances of 10 binary operations are carefully delineated, there is no description of the criteria for determining when usage is "complex and integrated" or "simple and nonintegrated." This variable would seem to be difficult to define. As older subjects could be assumed to have greater knowledge of the variables involved in a situation, the greater complexity of their statements might reflect their possession of more "bits" to integrate in addition to, or rather than, greater integrative ability.

2. The protocols of the 80% of subjects who immediately recognized the magnetic character of the apparatus are not compared with the protocols of the 20% who did not. For the former the problem was essentially one of explanation, while for the latter deductive problem solving was presumably necessary. Is the meaningful sample thus reduced to ten or eleven subjects principally from the youngest age group?

3. Tables of means, and other descriptive data, are missing from this paper. Relative frequencies of usage of the five forms, within group variations, and other important data are not reported.

4. The only statistical test reported is a $\chi^2 = .4917$, p greater than .05. We have no estimate of the probability of being wrong if we assume there were no differences.

5. Piagetian tasks are often difficult to reproduce. A complete description of what was done is necessary, particularly since the authors are refuting the procedure.

6. The experiment reported here does not suffice to refute the claim of Inhelder and Piaget that use of the complete lattice of operations is a necessary condition for the period of formal operations. The authors have instead demonstrated that this usage was not evidenced by a small sample of subjects on a particular task. However, they have not provided independent evidence that any of their subjects had fully attained the formal operational stage. Moreover, even if such evidence were presented, there would be no guarantee that a formal subject would use all of the operations available to him in any particular situation.

Suzanne K. Damarin
Richard J. Shumway
The Ohio State University